(19) JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 06332632 A
(43) Date of publication of application: 02.12.1994

(51) Int. Cl G06F 3/06
G06F 3/06, G06F 11/10

(21) Application number: 05125766
(22) Date of filing: 27.05.1993

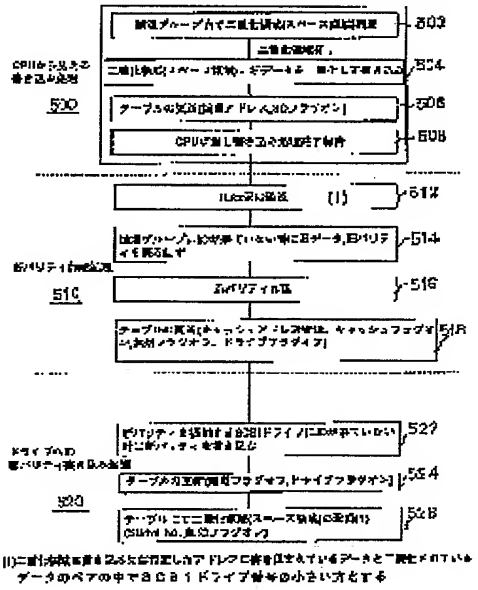
(71) Applicant: HITACHI LTD
(72) Inventor: TSUNODA HITOSHI
TAKAMOTO YOSHIFUMI

(54) DISK ARRAY DEVICE AND ITS CONTROL METHOD

(57) Abstract:

PURPOSE: To decrease the overhand at the time of writing data, in the disk array system of an RAID (level 5) for improving the processing performance by distributing the data.

CONSTITUTION: In a group for generating the parity, a duplex area is provided. At the time of write, data to be written in is written in the duplex area first (step 502), and at that time point, a write processing to a CPU is finished (508). The parity generation is executed at the subsequent suitable timing (516), and the data is written in a SCSI drive (522). That is, the generation of the parity is efficiently scheduled separately from write of the data. At the time of write of the data, the data to be written is written in the duplex area, and by processing independantly timewise write of the data to the SCSI drive and the generation of the parity, like this, the parity generation overhead at the time of write becomes invisible from a CPU.



(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号

特開平6－332632

(43)公開日　平成６年(1994)12月２日

(51)Int.Cl.[5]		識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０６Ｆ	3/06	３０５　Ｃ			
		３０１　Ｚ			
	11/10	３２０　Ｚ			

審査請求　未請求　請求項の数28　ＯＬ　（全 23 頁）

(21)出願番号　特願平5－125766

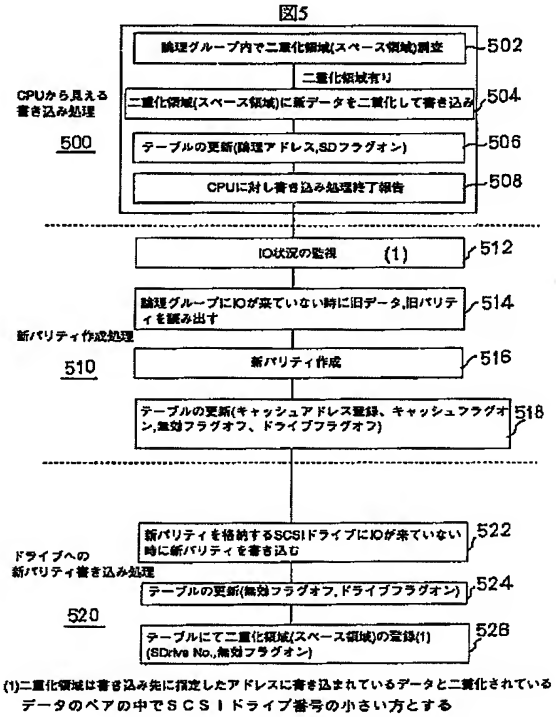
(22)出願日　平成５年(1993)５月27日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(72)発明者　角田　仁
東京都国分寺市東恋ケ窪１丁目280番地
株式会社日立製作所中央研究所内
(72)発明者　高本　良史
東京都国分寺市東恋ケ窪１丁目280番地
株式会社日立製作所中央研究所内
(74)代理人　弁理士　薄田　利幸

(54)【発明の名称】　ディスクアレイ装置及びその制御方法

(57)【要約】
【目的】データを分散させて処理性能を向上させるＲＡＩＤ（レベル５）のディスクアレイシステムにおいて、データの書込み時のオーバヘッドを減少させる。
【構成】パリティを作成されるグループにおいて、二重化領域を設ける。書き込み時において、書き込むデータを二重化領域にとりあえず書き込み（ステップ５０２）、その時点でＣＰＵに対し書き込み処理を終了とする（５０８）。パリティ作成は後の適当なタイミング（５１６）で行い、ＳＣＳＩドライブに書き込む（５２２）。つまり、パリティの作成をデータの書き込みとは別に、効率良くスケジューリングする。
【効果】データの書き込み時において、書き込むデータを二重化領域に書き込み、このように、データのＳＣＳＩドライブへの書き込みと、パリティの作成を、時間的に独立に処理することにより、書込み時のパリティ作成オーバヘッドはＣＰＵには見えなくなる。

図5



(1)二重化領域は書き込み先に指定したアドレスに書き込まれているデータと二重化されているデータのペアの中でＳＣＳＩドライブ番号の小さい方とする

1

【特許請求の範囲】

【請求項１】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブに格納する記憶装置の制御方法において、

データ書き込みに伴う誤り訂正符号の更新を、前記データ書き込み処理よりも遅延させて処理することを特徴とする記憶装置の制御方法。

【請求項２】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置におけるデータの格納方法であって、

前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置におけるデータの格納方法において、

データ書き込みに伴う誤り訂正符号の更新を、前記データ書き込み処理よりも遅延させて処理することを特徴とする記憶装置の制御方法。

【請求項３】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブに格納する記憶装置の制御方法において、

データ書き込みに伴う誤り訂正符号の更新を、データの書き込みを要求した上位置装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。

【請求項４】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置におけるデータの格納方法であって、

前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置におけるデータの格納方法において、

2

データ書き込みに伴う誤り訂正符号の更新を、データの書き込みを要求した上位置装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。

【請求項５】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブに格納するにおいて、

データ書き込みに伴う誤り訂正符号の更新を、前記制御装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。

【請求項６】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置におけるデータの格納方法であって、

前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置におけるデータの格納方法において、

データ書き込みに伴う誤り訂正符号の更新を、前記制御装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。

【請求項７】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブに格納するディスクアレイシステムにおいて、

データ書き込みに伴う誤り訂正符号の更新を、前記データの書き込みとは非同期でかつ所定のタイミング毎に処理することを特徴とするディスクアレイ装置の制御方法。

【請求項８】あるドライブに障害が発生し、障害が発生したドライブ内のデータまたはパリティを回復処理により復元し、この復元したデータまたはパリティを予備の書き込み領域に書き込んだ後、障害が発生したドライブを、正常なドライブに交換し、交換後は、この交換した正常なドライブは全てスペース領域により構成されているとして論理グループを再構成して処理を再開することを特徴とする請求項１記載の記憶装置の制御方法。

【請求項９】あるドライブに障害が発生したことを感知したら、障害が発生したドライブを正常なドライブに交

換し、障害が発生したドライブ内のデータまたはパリティを回復処理により復元し、この復元したデータまたはパリティと、障害が発生したドライブ内にあったスペース領域を、交換した正常なドライブに書き込んで再構成して処理を再開することを特徴とする請求項１記載の記憶装置の制御方法。

【請求項１０】最新に書き込まれたデータについては二重化して高信頼とし、書き込み要求があまり発行されないデータについてはパリティにより信頼性を確保するように、信頼性について二段階のレベルを設定することを特徴とする請求項１記載の記憶装置の制御方法。

【請求項１１】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置であって、

前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置において、

データ書き込みに伴う誤り訂正符号の更新を、前記データ書き込み処理よりも遅延させて処理することを特徴とする記憶装置。

【請求項１２】前記記憶装置を構成するドライブの集合の中に、書き込むデータを一旦二重化して書き込める領域をもつことを特徴とする請求項１１記載の記憶装置。

【請求項１３】スペース領域を、パリティを生成するデータと、それらのデータにより生成したパリティが格納されているドライブに分散してもつことを特徴とする請求項１１記載の記憶装置。

【請求項１４】スペース領域を、パリティおよびパリティの作成に関与したデータの格納されているドライブ以外の、異なる２台のドライブに確保することを特徴とする請求項１１記載の記憶装置。

【請求項１５】パリティを生成するデータと生成したパリティと、スペース領域が、それぞれ別のドライブに格納されており、これらの対応するデータ、パリティ、スペース領域が、各ドライブにおいて同一の物理的なアドレスに格納されるよう制御する手段を備えていることを特徴とする請求項１３記載の記憶装置。

【請求項１６】パリティを生成するデータと生成したパリティと、スペース領域が、それぞれ別のドライブに格納されており、これらの対応するデータ、パリティ、スペース領域は、各ドライブにおいて任意の物理的なアドレスに格納されるよう制御する手段を有することを特徴とする請求項１３記載の記憶装置。

【請求項１７】上位装置からの書込み要求に対し、ディスク装置を制御する制御装置が、書込みデータを、スペース領域の確保されている２台のドライブの物理的なアドレスに二重化して書き込む制御を行うプロセッサを持つことを特徴とする請求項１３記載の記憶装置。

【請求項１８】論理グループ内において、上位装置からの新データ書込み要求に対し、ディスク装置を制御する制御装置は、スペース領域の確保されている２台のドライブの物理的なアドレスに、二重化して書き込む様に変換して制御するプロセッサを持つことを特徴とする請求項１３記載の記憶装置。

【請求項１９】上位装置から書込み要求が発行された場合、書き込む新データをスペース領域の確保されている２台のドライブの物理的なアドレスに二重化して格納し、この段階で上位装置に対し、書き込み処理の終了を報告するプロセッサを持つことを特徴とする請求項１１記載の記憶装置。

【請求項２０】上位装置から書込み要求が発行された場合、書き込む新データをスペース領域の確保されている２台のドライブの物理的なアドレスに二重化して格納し、この段階で上位装置に対し、書き込み処理の終了を報告し、データの当該ドライブへの書き込み処理とは独立に、後で書き込みによるパリティの作成および当該ドライブへの書き込みを行うように制御するプロセッサを持つことを特徴とする請求項１８記載の記憶装置。

【請求項２１】上位装置から書込み要求が発行された場合、書き込む新データをスペース領域の確保されている２台のドライブの物理的なアドレスに二重化して格納し、この段階で上位装置に対し、書き込み処理の終了を報告した後、上位装置からの読み出しまたは書き込み要求数をカウントし、このカウント値をユーザまたはシステム管理者が予め設定した数とで比較する制御を行うプロセッサを持つことを特徴とする請求項１９記載の記憶装置。

【請求項２２】上位装置から書込み要求が発行された場合、書き込む新データをスペース領域の確保されている２台のドライブの物理的なアドレスに二重化して格納し、この段階で上位装置に対し、書き込み処理の終了を報告した後、上位装置からの読み出しまたは書き込み要求数をカウントし、このカウント値をユーザまたはシステム管理者が予め設定した数とで比較した結果、カウント値が設定数より小さく、しかも、当該ドライブに対し上位装置から読み出しまたは書き込み要求が発行されていない場合、書き込み処理におけるパリティの作成を開始するように判断する制御を行うプロセッサを持つことを特徴とする請求項２０記載の記憶装置。

【請求項２３】論理グループにおいて、パリティを作成するデータとパリティの集合をパリティグループとし、書き込み前と書き込み後では、パリティグループを構成するデータおよびパリティの格納されているドライブが異なることを特徴とする請求項１１記載の記憶装置。

【請求項２４】論理グループ内のあるドライブに障害が発生した場合、正常な残りのドライブ内のデータとパリティから、障害が発生したドライブ内のデータまたはパリティを回復処理により復元するが、この復元したデータまたはパリティをスペース領域に書き込む制御を行うプロセッサを持つことを特徴とする請求項１１記載のディスクアレイシステム。

【請求項２５】書き込みにより作成した新パリティを、当該ドライブへ書き込む前に、あるドライブに障害が発生した場合、旧パリティの作成に関与したデータについては、この旧パリティの作成に関与した、正常な残りのドライブ内のデータとパリティから回復処理により復元し、二重化されている新データが格納されているドライブに障害が発生した場合は、二重化データの一方から、障害が発生したドライブ内のデータまたはパリティを回復処理により復元する制御を行うプロセッサを持つことを特徴とする請求項１１記載のディスクアレイシステム。

【請求項２６】書き込みにより作成した新パリティを、当該ドライブへ書き込む前に、あるドライブに障害が発生した場合、正常な残りのドライブ内のデータとパリティと二重化されている新データから、障害が発生したドライブ内のデータまたはパリティを回復処理により復元するが、この復元したデータまたはパリティをスペース領域に書き込む制御を行うプロセッサを持つことを特徴とする請求項１１記載のディスクアレイシステム。

【請求項２７】データとパリティとスペース領域を、テーブルにより管理することを特徴とする請求項１１記載のディスクアレイシステム。

【請求項２８】論理グループ単位に、その内部にアドレス変換用テーブルと、パリティ生成回路と、キャッシュメモリと、それらを制御するマイクロプロセッサを持つことを特徴とする請求項１１記載のディスクアレイシステム。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明はコンピュータシステムに係り、特に高性能な入出力動作を可能とするディスクファイルシステムに関する。

【０００２】

【従来の技術】現在のコンピュータシステムにおいては、ＣＰＵ等の上位側が必要とするデータは２次記憶装置に格納され、ＣＰＵが必要とする時に応じ２次記憶装置に対してデータの書込み、読みだしを行っている。この２次記憶装置としては一般に不揮発な記憶媒体が使用され、代表的なものとして磁気ディスク装置（以下ドライブとする）、光ディスクなどがあげられる。

【０００３】近年高度情報化に伴い、コンピュータシステムにおいて、２次記憶装置の高性能化が要求されてきた。その一つの解として、多数の比較的容量の小さなドライブにより構成されるディスクアレイが考えられている。

【０００４】「D.Patterson,G.Gibson,and R.H.Kartz;A Case for Redundant Arrays of Inexpensive Disks(RAID),in ACM SIGMOD Conference,Chicago,IL,(June1988)」において、データを分割して並列に処理を行うディスクアレイ（レベル３）とデータを分散して、独立に扱うディスクアレイ（レベル５）について、その性能および信頼性の検討結果が報告されている。現在この論文に書かれている方式が最も一般的なディスクアレイと考えられている。

【０００５】以下にデータを分散して、独立に扱うディスクアレイ（レベル５）について説明する。レベル５のディスクアレイでは個々のデータを分割せずに独立に扱い、多数の比較的容量の小さなドライブに分散して格納するものである。現在、一般に使用されている汎用大型コンピュータシステムの２次記憶装置では、１ドライブ当りの容量が大きいため、他の読み出し／書込み要求に当該ドライブが使用されて、そのドライブを使用できずに待たされることが多く発生した。このタイプのディスクアレイでは汎用大型コンピュータシステムの２次記憶装置で使用されている大容量のドライブを、多数の比較的容量の小さなドライブで構成し、データを分散して格納してあるため、読み出し／書込み要求が増加してもディスクアレイの複数のドライブで分散して処理するため、読み出し／書込み要求がまたされることが減少する。しかし、ディスクアレイは、このように多数のドライブにより構成されるため、部品点数が増加し障害が発生する確率が高くなる。そこで、信頼性の向上を図る目的で、パリティを用意する必要がある。

【０００６】このパリティによりデータを格納したドライブに障害が発生した場合、その障害ドライブ内のデータを復元することが可能となる。ディスクアレイではデータからパリティを作成しデータと同様にドライブに格納しておく。この時、パリティは、パリティの作成に関与したデータとは別のドライブに格納される。

【０００７】これらのディスクアレイでは、現在一般に使用されている汎用大型コンピュータシステムと同様、２次記憶装置内では、個々のデータの格納場所（アドレス）は予め指定したアドレスに固定され、ＣＰＵから当該データへ読みだしまたは書込みする場合、この固定されたアドレスへアクセスすることになっている。この分散して格納するディスクアレイ（レベル５）ではストレージテクノロジコーポレーション（以下ＳＴＫとする）から製品発表がされている。米国特許ＷＯ　９１／２００７６では、レベル５の基本アーキテクチャにおいて、動的に変更可能なアドレスのテーブルを用意することにより、データ圧縮を行いデータの書込み処理において、トラック単位で書込み先のアドレスを動的に変換する方法について開示されている。

【０００８】また、特開平４－２３０５１２にはレベル５において、書き込み時に書き込むデータと、この書き込みにより更新したパリティを、それぞれ別の場所に書き込む方法について開示されている。さらに、ＩＢＭ社のディスクアレイ（９３３７）では、レベル５においてＷＡＤ（ライト　アシスト　デバイス）を設けることが発表されている。

【０００９】

【発明が解決しようとする課題】現在の汎用大型計算機システム等ではドライブにより構成される２次記憶装置内では、ＣＰＵから転送されてくるデータは個々のデータの格納場所（アドレス）が予め指定したアドレスに固定され、ＣＰＵから当該データへ読みだしまたは書込む場合は、この固定されたアドレスへアクセスすることになる。これは、ディスクアレイにおいても同じである。データを分割して並列に処理を行うディスクアレイ（レベル３）ではこのようにアドレスを固定しても影響は無いが、データを分散して、独立に扱うディスクアレイ（レベル５）ではアドレスを固定した場合、書込み時に大きな処理オーバヘッドが必要になる。以下それについて説明する。

【００１０】図１１は前記公知例でD.Pattersonらが提案したＲＡＩＤに述べられている、データを分散して、独立に扱うディスクアレイ（レベル５）内部のデータアドレスを示している。この各アドレスにあるデータは１回の読み出し／書込み処理される単位で、個々のデータは独立している。また、ＲＡＩＤで述べられているアーキテクチャではデータに対するアドレスは固定されている。前述したようにこのようなシステムでは、信頼性を向上するためパリティを設定することが不可欠である。本システムでは各ドライブ内の同一アドレスのデータによりパリティが作成される。すなわち、ドライブ＃１から４までのアドレス（１，１）のデータによりパリティが作成され、パリティを格納するドライブの（１，１）に格納される。本システムでは読み出し／書込み処理は現在の汎用大型計算機システムと同様に各ドライブに対し当該データをアクセスする。

【００１１】このようなディスクアレイにおいて、例えばドライブ＃３のアドレス（２，２）のデータを更新する場合、まず、更新される前のドライブ＃３の（２，２）のデータとパリティを格納してあるドライブの（２，２）のパリティを読みだし（１）、これらと更新する新しいデータとで排他的論理和をとり、新たなパリティを作成たする（２）。パリティの作成完了後、更新する新しいデータをドライブ＃３の（２，２）に、新パリティをパリティを格納するドライブの（２，２）に格納する（３）。

【００１２】図１３に示すように、このようなレベル５のディスクアレイでは、データの格納されているドライブ、パリティの格納されているドライブから古いデータとパリティを読みだすため、ディスクを平均１／２回転待ち、それから読みだしてパリティを作成する。この新しく作成したパリティを書き込むため更に一回転必要となり、データを書き替える場合最低で１．５回転待たなければならない。ドライブにおいては１．５回転ディスクの回転を待つということは非常に大きなオーバヘッドとなる。このような書込み時のオーバヘッドを削減するため、書込み先のアドレスを動的に変換する方法が考えられ、ＷＯ　９１／２００７６に開示されている。

【００１３】また、前記特開平４－２３０５１２においても、書き込み時において書き込みデータをそのまま、書き込みデータが書き込まれるアドレスではなく別のアドレスに、書き込むことにより書き込みオーバヘッドを削減する方法について開示されている。ＣＰＵ側から書き込むデータが送られてくるとですぐにパリティの更新を行ない、更新後のパリティを書き込む。このように、レベル５のディスクアレイでは、読みだしと比較し書き込み時ではパリティ生成とこの生成したパリティを書き込む処理のオーバヘッドが非常に大きいため、ＣＰＵからの読みだし、書き込み要求が多いときには、この処理オーバヘッドが性能低下の大きな原因となる。

【００１４】本発明の目的は、レベル５のディスクアレイ装置において、書き込み時における処理のオーバヘッドを減少させて、ディスクアレイ装置の性能向上を図ることにある。

【００１５】本発明の他の目的は、障害ドライブ内のデータ復元用スペアドライブを装置の性能向上に利用することによってドライブ資源の有効活用を図ることにある。

【００１６】

【課題を解決するための手段】本発明ではパリティグループを構成するドライブと二重化領域（スペース領域）のドライブにより論理グループを構成し、このスペース領域を有効に活用することにより、高信頼性を保ちながら、しかも、書き込み時のパリティ更新の開始時間を遅らせ、後のＣＰＵからの読みだしまたは書き込み要求が少ないときにパリティ生成を行う。

【００１７】具体的には、書き込み時に、論理グループ１０を構成するＳＣＳＩドライブ１２の中で、書き込むデータ（新データ）をとりあえずスペース領域に二重化して格納する。ＣＰＵ１に対してはこの時点で書き込み処理を完了したと報告する。

【００１８】また、パリティの作成および当該ＳＣＳＩドライブ１２へのパリティの書き込みは、新データのＳＣＳＩドライブへの書き込みとは独立のタイミングで、処理する。具体的には、ＡＤＣ２のＭＰ１　２０が当該論理グループ１０に対するＣＰＵ１からの読みだし／書き込み要求をカウントし、予めユーザまたはシステム管理者が設定した数より少ない場合で、しかも当該ＳＣＳＩドライブ１２に対し読みだしまたは書き込み要求が発

行されていないときにパリティの作成を行い、パリティの作成完了後当該ＳＣＳＩドライブ１２に対しパリティを書き込む。

【００１９】パリティの書き込みの他の方法として、一定時間毎の割込み処理で行なってもよい。一日の中でＣＰＵからの読みだしまたは書き込み要求数の少ない時間帯、あるいは一月の中で少ない日を予測し、スケジュール化しておけばよい。

【００２０】パリティの作成およびそのパリティの当該ＳＣＳＩドライブ１２へのパリティの書き込みが完了する前に、当該論理グループ１０において任意の１台のＳＣＳＩドライブに対し障害が発生し、その内部のデータが読み出せなくなった場合は、二重化データ以外のデータが格納されているＳＣＳＩドライブ１２に対しては、前のパリティと、残っているデータから障害が発生したＳＣＳＩドライブ１２内のデータを回復することが可能で、二重化されている新しいデータにおいては、障害が発生していない方のＳＣＳＩドライブ１２内の新データにより回復することが可能である。

【００２１】

【作用】本発明では、上記のようにデータの書き込みとパリティの作成およびＳＣＳＩドライブ１２への書き込みを独立させることにより、ユーザ（ＣＰＵ）からは書き込み時のパリティ作成によるオーバヘッドはなくなる。これは、ＣＰＵからの読みだしまたは書き込み要求数には時間的変動があるため、読みだしまたは書き込み要求数が多いときには、書き込み処理におけるパリティの更新をその都度行わず、データの書き込みが完了した時点でＣＰＵには終了を報告し、比較的読みだしまたは書き込み要求の数が少ないときまでパリティの更新を遅らせる。このパリティの更新はＣＰＵ側の関知はなく、ディスクアレイコントローラ２が独自に行う。

【００２２】このため、ＣＰＵ側から見たとき、従来のディスクアレイ（ＲＡＩＤ方式）では図１２に示すように書き込み時に平均１．５回転の回転待ち時間を必要としたのが、本発明によれば平均０．５回転の回転待ち時間ですむ。また、信頼性の面から見ても従来のディスクアレイ（ＲＡＩＤ方式）と比較し、同等に向上させることが可能となる。

【００２３】

【実施例】

〈実施例１〉以下本発明の一実施例を、図１～図５及び図１３により説明する。図１において、本実施例はＣＰＵ１、ディスクアレイコントローラ（以下ＡＤＣ）２，ディスクアレイユニット（以下ＡＤＵ）３により構成される。ＡＤＵ３は複数の論理グループ１０により構成され，個々の論理グループ１０はm台のＳＣＳＩドライブ１２と、各々のＳＣＳＩドライブ１２とＡＤＣ２を接続するドライブパス９－１から４により構成される。なお，このＳＣＳＩドライブ１２の数は本発明の効果を得るには、特に制限は無い。この論理グループ１０は障害回復単位で、この論理グループ１０内の各ＳＣＳＩドライブ１２内の各データによりパリティを作成する。本実施例ではm－１台の個々のＳＣＳＩドライブ１２内のデータから各々のパリティが作成される。

【００２４】次にＡＤＣ２の内部構造について図１を用いて説明する。ＡＤＣ２はチャネルパスディレクタ５と２個のクラスタ１３とバッテリバックアップ等により不揮発化された半導体メモリであるキャッシュメモリ７により構成される。このキャッシュメモリ７にはデータとアドレス変換用テーブルが格納されている。このキャッシュメモリ７およびその中のアドレス変換用テーブルはＡＤＣ２内の全てのクラスタにおいて共有で使用される。クラスタ１３はＡＤＣ２内において独立に動作可能なパスの集合で、各クラスタ１３間においては電源、回路は全く独立となっている。クラスタ１３はチャネル、キャッシュメモリ７間のパスである、チャネルパス６と、キャッシュメモリ７、ＳＣＳＩドライブ１２間のパスであるドライブパス６－１から４が、それぞれ、２個ずつで構成されている。それぞれのチャネルパス６－１から４とドライブパス８はキャッシュメモリ７を介して接続されている。ＣＰＵ１より発行されたコマンドは外部インターフェースパス４を通ってＡＤＣ２のチャネルパスディレクタ５に発行される。ＡＤＣ２は２個のクラスタ１３により構成され、それぞれのクラスタは２個のパスで構成されるため、ＡＤＣ２は合計４個のパスにより構成される。このことから、ＡＤＣ２ではＣＰＵ１からのコマンドを同時に４個まで受け付けることが可能である。そこで、ＣＰＵ１からコマンドが発行された場合ＡＤＣ２内のチャネルパスディレクタ５によりコマンドの受付が可能かどうか判断する。

【００２５】図２は図１のチャネルパスディレクタ５と１クラスタ１３－１内の内部構造を示した図である。図２に示すように、ＣＰＵ１からＡＤＣ２に送られてきたコマンドはインターフェースアダプタ（以下ＩＦ　Ａｄｐ）１５により取り込まれ、マイクロプロセッサであるＭＰ１　２０はクラスタ内の外部インターフェースパス４の中で使用可能なパスがあるかを調べ、使用可能な外部インターフェースパス４がある場合はＭＰ１　２０はチャネルパススイッチ１６を切り換えてコマンドの受付け処理を行ない、受け付けられない場合は受付不可の応答をＣＰＵ１へ送る。

【００２６】本実施例ではＡＤＵ３を構成するＳＣＳＩドライブ１２はＳＣＳＩインターフェースのドライブを使用する。ＣＰＵ１をＩＢＭシステム９０００シリーズのような大型汎用計算機とした場合、ＣＰＵ１からはＩＢＭオペレーティングシステム（ＯＳ）で動作可能なチャネルインターフェースのコマンド体系にのっとってコマンドが発行される。そこで、ＳＣＳＩドライブ１２をＳＣＳＩインターフェースのドライブを使用した場合、

ＣＰＵ１からのコマンドを、ＳＣＳＩインターフェースのコマンド体系にのっとったコマンドに変換する必要が生じる。この変換はコマンドのプロトコル変換と、アドレス変換に大きく分けられる。以下にアドレス変換について説明する。

【００２７】ＣＰＵ１から指定されるアドレスは、図１２に示すように当該データが格納されているトラックが所属するシリンダの位置と、そのシリンダ内において当該データが格納されているトラックを決定するヘッドアドレスと、そのトラック内のレコードの位置を特定する。具体的には要求データが格納されている当該ドライブの番号（ＣＰＵ指定ドライブ番号）と当該ドライブ内のシリンダ番号であるシリンダアドレス（ＣＣ）とシリンダにおいてトラックを選択するヘッドの番号であるヘッドアドレス（ＨＨ）とレコードアドレス（Ｒ）からなるＣＣＨＨＲである。

【００２８】従来のＣＫＤフォーマット対応の磁気ディスクサブシステム（ＩＢＭ３９９０－３３９０）ではこのアドレスに従ってドライブへアクセスすれば良い。しかし、本実施例では複数のＳＣＳＩドライブ１２により従来のＣＫＤフォーマット対応の磁気ディスクサブシステムを論理的にエミュレートする。つまり、ＡＤＣ２は複数のＳＣＳＩドライブ１２が、従来のＣＫＤフォーマット対応の磁気ディスクサブシステムで使用されているドライブ１台に相当するようにＣＰＵ１にみせかける。このため、ＣＰＵ１から指定してきたアドレス（ＣＣＨＨＲ）をＳＣＳＩドライブのアドレスにＭＰ１　２０は変換する。このアドレス変換には以下に示すようなアドレス変換用のテーブル４０（以下アドレステーブルとする）が使用される。

【００２９】ＡＤＣ２内のキャッシュメモリ７には、その内部の適当な領域に図３に示すようなアドレステーブル４０が格納されている。本実施例では、ＣＰＵ１が指定してくるドライブはＣＫＤフォーマット対応の単体ドライブである。しかし、本発明ではＣＰＵ１は単体と認識しているドライブが、実際は複数のＳＣＳＩドライブ１２により構成されるため、論理的なドライブとして定義される。このため、ＡＤＣ２のＭＰ１　２０はＣＰＵ１より指定してきたアドレス（ＣＰＵ指定ドライブ番号４１とＣＣＨＨＲ４６）をＳＣＳＩドライブ１２に対するＳＣＳＩドライブアドレス４２（ＳＣＳＩドライブ番号４３とそのＳＣＳＩドライブ内のアドレス（以下ＳＣＳＩ内Ａｄｄｒとする）４４）に変換する。

【００３０】アドレステーブル４０はＣＰＵ１が指定するＣＰＵ指定ドライブ番号４１とＳＣＳＩドライブアドレス４２により構成される。ＳＣＳＩドライブアドレス４２はＳＣＳＩドライブ１２のアドレスであるＳＣＳＩドライブ番号４３とそのＳＣＳＩドライブ内の実際にデータが格納されているアドレスである、ＳＣＳＩ内Ａｄｄｒ４４と、論理グループ１０内において、そのＳＣＳＩ内Ａｄｄｒ４４により決定されるパリティグループにおけるパリティが格納されているＳＣＳＩドライブ番号（パリティドライブ番号５０）と、二重化領域（スペース領域）が格納されているＳＣＳＩドライブ番号（スペースドライブ番号５１）により構成されている。このアドレステーブル４０では、論理アドレス４５によりＳＣＳＩドライブ番号４３とＳＣＳＩ内Ａｄｄｒ４４を決定する。アドレステーブル４０のＳＣＳＩドライブアドレス４２に登録されているＳＣＳＩドライブ番号４３のＳＣＳＩドライブ１２により論理グループ１０は構成される。

【００３１】この論理グループ１０内の同一ＳＣＳＩ内Ａｄｄｒ４４において、パリティが格納されているＳＣＳＩドライブ番号４３をパリティドライブ番号５０に登録し、スペース領域が確保されているＳＣＳＩドライブ番号４３をスペースドライブ番号５１に登録する。スペースドライブ番号５１にはスペース領域が確保されているＳＣＳＩドライブ番号の他にＳＤフラグ５３により構成される。ＳＤフラグ５３は、スペース領域に格納されているデータが有効で、書き込み処理においてスペース領域として使用できない場合はオン（１）となり、無効で、スペース領域として使用可能な場合はオフ（０）となる。この論理グループ１０はデータとこのデータと関連するパリティにより構成されるパリティグループとスペース領域により構成される。

【００３２】論理アドレス４５にはＣＰＵ１から指定されるアドレスである、ＣＰＵ指定ドライブ番号４１とＣＣＨＨＲの中でＣＣＨＨＲ４６が登録されており、それ以外にはこの論理アドレス４５のデータがキャッシュメモリ７内に存在する場合の、そのデータのキャッシュメモリ７内のアドレスを格納するキャッシュアドレス４７と、キャッシュメモリ７内にその論理アドレス４５のデータを保持している場合オン（１）が登録されるキャッシュフラグ４８と、その論理アドレス４５にはスペース領域が確保されている場合オン（１）となる無効フラグ４９とキャッシュメモリ７内の書き込みデータがドライブに書き込まれている場合オン（１）となるドライブフラグ５２により構成される。

【００３３】以上のようにアドレステーブル４０によりＣＰＵ指定ドライブ番号４１とＣＣＨＨＲ４６を論理アドレス４５に変換し、そのデータが実際に格納されているＳＣＳＩドライブ番号４３とＳＣＳＩ内Ａｄｄｒ４４を決定する。

【００３４】例えば、図３においてＣＰＵ１からＣＰＵ指定ドライブ番号４１としてＤｒｉｖｅ＃１、ＣＣＨＨＲ４６がＡＤＲ８のデータに対し要求を発行してきた場合、アドレステーブル４０においてＣＰＵ指定ドライブ番号４１がＤｒｉｖｅ＃１の領域の各論理アドレス４５内のＣＣＨＨＲ４６を調べ、ＣＣＨＨＲ４６がＡＤＲ８の論理アドレス４５を探す。図３においては論理アドレ

ス４５としてＤａｔａ＃２３（Ｄ＃２３）がＣＣＨＨＲ４６がＡＤＲ８となっており、Ｄａｔａ＃２３（Ｄ＃２３）が当該論理アドレス４５である。

【００３５】このＤａｔａ＃２３（Ｄ＃２３）はアドレステーブル４０からＳＤ＃２のＳＣＳＩインターフェースのＳＣＳＩドライブ１２内のＳＣＳＩ内Ａｄｄｒ４４としてＤＡＤＲ８に該当することが分かり、物理的なアドレスへ変換される。また、このＤａｔａ＃２３（Ｄ＃２３）に対応するパリティは、パリティドライブ番号５０からＤａｔａ＃２３（Ｄ＃２３）と同一のＳＣＳＩ内Ａｄｄｒ４４のＤＡＤＲ８のＳＤ＃４のＳＣＳＩドライブ１２に格納されており、スペアドライブ番号５１から、ＳＤフラグ５３がオン（１）のため、ＳＤ＃４、５のＳＣＳＩ内Ａｄｄｒ４４がＤＡＤＲ８に二重化して格納されているデータは有効で、この領域は二重化領域（スペア領域）として使用することは禁止されている。

【００３６】このように、ＣＰＵ１から指定されたアドレスを論理アドレス４５に変換し、実際に読みだし／書き込みを行うＳＣＳＩドライブ１２の物理的なアドレスに変換した後、ＳＤ＃２のＳＣＳＩドライブ１２のＤａｔａ＃２３（Ｄ＃２３）に対し読み出しまたは書込み要求が発行される。この時アドレステーブル４０においてＤａｔａ＃２３（Ｄ＃２３）の論理アドレス４５ではキャッシュフラグ４８がオン（１）のためこのデータはキャッシュメモリ７内のＣＡＤＲ２，１に存在する。もし、キャッシュフラグ４８がオフ（０）であればＣＡＤＲ２，４のキャッシュメモリ７内には、当該データは存在しない。また、このデータは無効フラグ４９がオフ（０）のため、このデータは有効となる。さらにドライブフラグ５２がオン（１）のため、このデータはキャッシュメモリ７からドライブに既に書き込まれている。

【００３７】この、アドレステーブル４０はシステムの電源をオンした時に、ＭＰ１　２０により論理グループ１０内のある特定のＳＣＳＩドライブ１２から、キャッシュメモリ７にＣＰＵ１は関知せず自動的に読み込まれる。一方、電源をオフする時はＭＰ１　２０によりキャッシュメモリ７内のアドレステーブル４０を、読み込んできたＳＣＳＩドライブ１２内の所定の場所にＣＰＵ１は関知せずに自動的に格納する。

【００３８】次に、ＡＤＣ２内での具体的なＩ／Ｏ処理について図１、図２を用いて説明する。ＣＰＵ１より発行されたコマンドはＩＦ　Ａｄｐ１５を介してＡＤＣ２に取り込まれ、ＭＰ１　２０により読み出し要求か書込み要求か解読される。まず、読み出し要求の場合の処理方法を以下に示す。

【００３９】ＭＰ１　２０が読み出し要求のコマンドを認識すると、ＭＰ１　２０はＣＰＵ１から送られてきたＣＰＵ指定ドライブ番号４１とＣＣＨＨＲ４６（以下両方を併せてＣＰＵ指定アドレスとする）についてアドレステーブル４０を参照し、当該データの論理アドレス４５への変換を行ない、論理アドレス４５に登録されているキャッシュメモリ７内に存在するかどうかキャッシュフラグ４８を調べ、判定する。

【００４０】キャッシュフラグ４８がオンでキャッシュメモリ７内に格納されている場合（キャッシュヒット）は、ＭＰ１　２０がキャッシュメモリ７から当該データを読みだす制御を開始し，キャッシュメモリ７内に無い場合（キャッシュミス）は当該ドライブ１２へその内部の当該データを読みだす制御を開始する。

【００４１】キャッシュヒット時、ＭＰ１　２０はアドレステーブル４０によりＣＰＵ１から指定してきたＣＰＵ指定アドレスから論理アドレス４５に変換し、論理アドレス４５内のキャッシュアドレス４７によりキャッシュメモリ７のアドレスに変換し、キャッシュメモリ７へ当該データを読み出しに行く。具体的にはＭＰ１　２０の指示の元でキャッシュアダプタ回路（Ｃ　Ａｄｐ）２４によりキャッシュメモリ７から当該データは読み出される。

【００４２】Ｃ　Ａｄｐ２４はキャッシュメモリ７に対するデータの読みだし、書き込みをＭＰ１　２０の指示で行う回路で、キャッシュメモリ７の状態の監視、各読みだし、書き込み要求に対し排他制御を行う回路である。Ｃ　Ａｄｐ２４により読み出されたデータはデータ制御回路（ＤＣＣ）２２の制御によりチャネルインターフェース回路（ＣＨ　ＩＦ）２１に転送される。ＣＨ　ＩＦ２１ではＣＰＵ１におけるチャネルインターフェースのプロトコルに変換し、チャネルインターフェースに対応する速度に速度調整する。具体的にはＣＰＵ１，ＡＤＣ２間のチャネルインターフェースを光のインターフェースにした場合、光のインターフェースのプロトコルをＡＤＣ２内では電気処理でのプロトコルに変換する。ＣＨ　ＩＦ２１におけるプロトコル変換および速度調整後は、チャネルパスディレクタ５において、チャネルパススイッチ１６が外部インターフェースパス４を選択しＩＦＡｄｐ１５によりＣＰＵ１へデータ転送を行なう。

【００４３】一方、キャッシュミス時はキャッシュヒット時と同様にアドレステーブル４０により、ＣＰＵ指定アドレスを論理アドレス４５に変換し、論理アドレス４５から当該ＳＣＳＩドライブ番号とそのＳＣＳＩドライブ内の実際にデータが格納されているＳＣＳＩ内Ａｄｄｒ４４を認識し、そのアドレスに対し、ＭＰ１　２０はＤｒｉｖｅ　ＩＦ２８に対し、当該ドライブ１２への読み出し要求を発行するように指示する。Ｄｒｉｖｅ　ＩＦ２８ではＳＣＳＩの読み出し処理手順に従って、読み出しコマンドをドライブユニットパス９－１または９－２を介して発行する。Ｄｒｉｖｅ　ＩＦ２８から読み出しコマンドを発行された当該ＳＣＳＩドライブ１２においては指示されたＳＣＳＩ内Ａｄｄｒ４４へシーク、回転待ちのアクセス処理を行なう。当該ＳＣＳＩドライブ１２におけるアクセス処理が完了した後、当該ＳＣＳＩ

ドライブ１２は当該データを読み出しドライブユニットパス９を介してＤｒｉｖｅ　ＩＦ２８へ転送する。

【００４４】Ｄｒｉｖｅ　ＩＦ２８では転送されてきた当該データをＳＣＳＩドライブ側のキャッシュアダプタ回路（Ｃ　Ａｄｐ）１４に転送し、（Ｃ　Ａｄｐ）１４ではキャッシュメモリ７にデータを格納する。この時、Ｃ　Ａｄｐ１４はキャッシュメモリ７にデータを格納することをＭＰ１　２０に報告し、ＭＰ１　２０はこの報告を元にアドレステーブル４０のＣＰＵが読みだし要求を発行したＣＰＵ指定アドレスに対応した論理アドレス４５のキャッシュフラグ４８をオン（１）にし、キャッシュアドレス４７にキャッシュメモリ７内のデータを格納したアドレスを登録する。キャッシュメモリ７にデータを格納し、アドレステーブル４０のキャッシュフラグ４８をオン（１）にし、キャッシュメモリ７内のアドレスを更新した後はキャッシュヒット時と同様な手順でＣＰＵ１へ当該データを転送する。

【００４５】一方書き込み時は以下のように処理される。書き込み処理はユーザが書き込み先のアドレス（ＣＰＵ指定アドレス）を指定し、その位置にユーザはデータを書き込んでいると認識する。つまりユーザは固定の位置にアドレスを指定していると認識している。

【００４６】ＣＰＵ１から、指定アドレスすなわち、図３に示すアドレステーブル４０においてＣＰＵ指定ドライブ番号４１がドライブ＃１でＣＣＨＨＲ４６がＡＤＲ８に対して、書き込み命令が発行されたとする。まず、ＡＤＣ２のＭＰ１　２０はＣＰＵ１からドライブ＃１のＣＣＨＨＲ４６がＡＤＲ８に対し書込み要求のコマンドを受け取った後、コマンドを受け取ったＭＰ１　２０が所属するクラスタ１３内の各チャネルパス６において処理可能かどうかを調べ、可能な場合は処理可能だという応答をＣＰＵ１へ返す。ＣＰＵ１では処理可能だという応答を受け取った後にＡＤＣ２へデータを転送する。この時、ＡＤＣ２ではＭＰ１　２０の指示によりチャネルパスディレクタ５において、チャネルパススイッチ１６が当該外部インターフェースパス４とＩＦ　Ａｄｐ１５を当該チャネルパス６と接続しＣＰＵ１とＡＤＣ２間の接続を確立する。

【００４７】ＣＰＵ１とＡＤＣ２間の接続を確立後ＣＰＵ１からのデータ転送を受け付ける。ＣＰＵ１から転送されてきた書き込みデータ（以下新データとする）はＭＰ１２０の指示により、ＣＨ　ＩＦ２１によりプロトコル変換を行ない、外部インターフェースパス４での転送速度からＡＤＣ２内での処理速度に速度調整する。ＣＨ　ＩＦ２１におけるプロトコル変換および速度制御の完了後、データはＤＣＣ２２によるデータ転送制御を受け、Ｃ　Ａｄｐ２４に転送され、Ｃ　Ａｄｐ２４によりキャッシュメモリ７内に格納される。

【００４８】この時、ＣＰＵ１から送られてきた情報が、ＣＰＵ指定アドレスの場合は、データが転送されてくる前に必ずＣＰＵ指定アドレスがＣＰＵ１より転送されているため、読みだしと同様にアドレステーブル４０によりアドレス変換を行い、論理アドレス４５に変換する。また、ＣＰＵ１から送られてきた情報がデータの場合は、キャッシュメモリ７に格納したアドレスを上記アドレス変換により変換した論理アドレス４５のキャッシュアドレス４７に登録する。この時、書き込みデータをキャッシュメモリ７内に保持するときは、論理アドレス４５のキャッシュフラグ４８をオン（１）とし、保持しない場合はキャッシュフラグ４８をオフ（０）とする。

【００４９】なお、キャッシュメモリ７内に保持されている新データに対し、さらに書き込み要求がＣＰＵ１から発行された場合は、キャッシュメモリ７内に保持されている新データを書き替える。

【００５０】キャッシュメモリ７に格納された新データは、この新データにより新しくパリティを更新し（以下更新されたパリティを新パリティとする）、以下のように論理グループ１０内のＳＣＳＩドライブ１２へ新データと新パリティを格納する。

【００５１】図３に示すようにスペース領域及びパリティはデータと同じ様に扱われ、論理グループ１０内の各ＳＣＳＩドライブ１２に分散して格納されている。論理グループ１０を構成する各ＳＣＳＩドライブ１２内のデータについては行方向（同一ＳＣＳＩ内Ａｄｄｒ４４）にパリティグループが構成され、このパリティグループ内のデータにおいてパリティが作成される。つまり、パリティグループはデータとパリティにより構成され、論理グループ１０はパリティグループとスペース領域で構成される。

【００５２】図３において具体的な例を示す。ＳＣＳＩ内Ａｄｄｒ４４がＤＡＤＲ１については、ＳＤ＃１のＳＣＳＩドライブ１２に格納されているＤａｔａ＃１（Ｄ＃１）と、ＳＤ＃２のＳＣＳＩドライブ１２に格納されているＤａｔａ＃２（Ｄ＃２）と、ＳＤ＃３のＳＣＳＩドライブ１２に格納されているＤａｔａ＃３（Ｄ＃３）によりパリティが作成される。このパリティがＳＤ＃６のＳＣＳＩドライブ１２に格納され、これらがパリティグループを構成する。また、論理グループ１０は、ＳＤ＃４のＳＣＳＩドライブ１２に確保されているスペース領域（Ｓ）と、ＳＤ＃５のＳＣＳＩドライブ１２に確保されているスペース領域（Ｓ）と、上記パリティグループにより構成される。

【００５３】ＭＰ１　２０はアドレステーブル４０を参照し、データ、スペース領域、パリティが格納されているＳＣＳＩドライブ１２を認識する。具体的には、アドレステーブル４０においてＣＰＵ指定アドレスが指定したＣＰＵ指定ドライブ番号４１に対応する領域において、ＣＰＵ指定アドレスで指定したＣＣＨＨＲ４６と一致しているＳＣＳＩドライブアドレス４２に登録されている論理アドレス４５をＭＰ１　２０は探す。ＭＰ１

２０はＣＰＵ指定アドレスから論理アドレス４５への変換後、論理アドレス４５からその論理アドレス４５が格納されているＳＣＳＩドライブ番号４３と、そのＳＣＳＩドライブ１２内の物理的なアドレスであるＳＣＳＩ内Ａｄｄｒ４４に変換する。

【００５４】パリティは、論理グループ１０を構成する各ＳＣＳＩドライブ１２において、各ＳＣＳＩ内Ａｄｄｒ４４が同一のデータに対し作成され、そのパリティもデータと同一のＳＣＳＩ内Ａｄｄｒ４４に格納される。このため、アドレステーブル４０のパリティドライブ番号５０とスペースドライブ番号５１において、パリティおよびスペース領域はそれぞれが格納されているＳＣＳＩドライブ番号４３のみが登録されている。そこで、ＭＰ１　２０ではアドレステーブル４０によりパリティドライブ番号５０とスペースドライブ番号５１を決定することが可能となる。つまり、パリティドライブ番号５０とスペースドライブ番号５１を決定することにより、パリティおよびスペース領域はデータと同一のＳＣＳＩ内Ａｄｄｒ４４に格納されていることから、それらのアドレスを決定することになる。このように、データ、スペース領域、パリティが格納されているＳＣＳＩドライブ１２を認識した後は、Ｄｒｉｖｅ　ＩＦ２８に対し、各々の当該ＳＣＳＩドライブ１２に対し書き込み処理を行なうように指示する。

【００５５】本発明における書き込み処理とは、論理グループ１０において新データをキャッシュメモリ７から実際にＳＣＳＩドライブ１２に書き込む処理と、新データの書き込みによりパリティを新たに作り直さなければならないため、この新パリティを作成するための書き込み前のデータ（以下旧データ）と書き込み前のパリティ（以下旧パリティとする）を読み出し、新パリティを作成する処理と、書き込み後の新パリティを実際にＳＣＳＩドライブ１２に書き込む一連の処理をいう。この新データをキャッシュメモリ７に格納した後の一連の書き込み処理のフローチャートを図５に示す。

【００５６】図４に示すようにＣＰＵ１からＳＤ＃１のＳＣＳＩドライブ１２のＤａｔａ＃１（Ｄ＃１）の論理アドレスに対し、ＮＤ＃１のデータの書き込み要求が発行された場合、先に示したように、ＣＰＵ１からキャッシュメモリ７に新データは一旦格納される。新データのキャッシュメモリ７への格納後、書き込み処理は次の手順で行っていく。キャッシュメモリ７に新データ（ＮＤ＃１）が格納された後、ＭＰ１　２０は書き込み処理に入り、アドレステーブル４０によりＳＤ＃１のＳＣＳＩドライブ１２が所属する論理グループ１０で、Ｄａｔａ＃１（Ｄ＃１）のＳＣＳＩ内Ａｄｄｒ４４であるＤＡＤＲ１に対し、スペース領域が確保されているＳＤ＃４，５のＳＣＳＩドライブ１２に対し、使用権の獲得を行なう。

【００５７】ＳＤ＃４、５のＳＣＳＩドライブ１２の使用権を獲得した後は図５のフローチャートに示すように書き込み処理５００を行なっていく。まず、ＭＰ１　２０はアドレステーブル４０のスペースドライブ番号５１のＳＤフラグ５３をチェックし、ＳＤフラグ５３がオフ（０）の場合、スペース領域として使用でき、オン（１）の場合は使用できないと判断する（５０２）。ＭＰ１　２０は、このＳＤフラグ５３によりＳＤ＃４，５のＳＣＳＩドライブ１２にスペース領域が確保されているかを判断する。そして、ＳＤフラグ５３がオフ（０）の場合、キャッシュメモリ７に格納されているＮＤ＃１をこのＳＤ＃４，５に二重化して書き込み（５０４）、ＭＰ１　２０はＣＰＵ１へ書き込み完了の報告を行う（５０８）。

【００５８】一方、ステップ５０２のチェックにより、ＳＤフラグ５３がオン（１）の場合は、図１３に示すように、書き込む新データ（ＮＤ＃１）をキャッシュメモリ７に書き込んだ後、ＭＰ１　２０は優先的に前の書き込み処理におけるパリティの作成を指示し、このパリティを当該ＳＣＳＩドライブ１２に書き込む（１３１０）。前の書き込み処理におけるパリティの作成が完了し、当該ＳＣＳＩドライブへの書き込みが完了するとＭＰ１　２０は、アドレステーブル４０においてスペースドライブ番号５１にＳＤフラグ５３をオフ（０）とし（１３０８）、キャッシュメモリ７に格納されているＮＤ＃１を二重化して書き込み（１３１６）、ＭＰ１　２０はＣＰＵ１へ書き込み完了の報告を行う（１３２０）。

【００５９】そこで、次にＳＣＳＩドライブ１２への新データ（ＮＤ＃１）の書き込み方法を示す。ＭＰ１　２０はアドレステーブル４０のＳＤフラグ５３がオフ（０）になっているのを確認後、Ｄｒｉｖｅ　ＩＦ２８に書き込むデータである新データ（ＮＤ＃１）をスペース領域が確保されているＳＤ＃４とＳＤ＃５のＳＣＳＩドライブ１２に書き込むよう指示する。Ｄｒｉｖｅ　ＩＦ２８ではＳＣＳＩの書込み手順にしたがってドライブユニットパス９－１から４の中の２本を介してＳＤ＃４とＳＤ＃５のＳＣＳＩドライブ１２に書込みコマンドを発行する。

【００６０】Ｄｒｉｖｅ　ＩＦ２８から書込みコマンドを発行された当該ＳＣＳＩドライブ１２では、Ｄｒｉｖｅ　ＩＦ２８から送られてきたＣＰＵ１が書き込み先のアドレスとして指定したＣＰＵ指定アドレスを論理アドレス４５に変換し、論理アドレス４５であるＤａｔａ＃１に対応するＳＣＳＩ内Ａｄｄｒ４４のＤＡＤＲ１へシーク、回転待ちのアクセス処理を行なう。ＳＤ＃４、５のＳＣＳＩドライブ１２においてアクセス処理が完了し書込みが可能になり次第、Ｃ　Ａｄｐ１４はキャッシュメモリ７から書き込む新データ（ＮＤ＃１）を読み出してＤｒｉｖｅＩＦ２８へ転送し、Ｄｒｉｖｅ　ＩＦ２８では転送されてきた新データ（ＮＤ＃１）をドライブユ

ニットパス９－１から４の中の２本を介してＳＤ＃４，５のＳＣＳＩドライブ１２へ転送する。新データ（ＮＤ＃１）のＳＤ＃４，５のＳＣＳＩドライブ１２への書込みが完了すると、ＳＤ＃４，５のＳＣＳＩドライブ１２はＤｒｉｖｅ　ＩＦ２８に完了報告を行ない、Ｄｒｉｖｅ　ＩＦ２８がこの完了報告を受け取ったことを、ＭＰ１　２０に報告する。

【００６１】この時、アドレステーブル４０において、書き込み前の旧データの論理アドレス４５（Ｄａｔａ＃１，Ｄ＃１）の無効フラグをオン（１）とし、書き込み前の論理アドレス４５（Ｄａｔａ＃１，Ｄ＃１）内のＣＨＨＲ４６のアドレスを、新データ（ＮＤ＃１）を二重化して書き込んだスペース領域の２つの論理アドレス４５のＣＣＨＨＲ４６に登録し、無効フラグをオフ（０）とし、ドライブフラグ５２をオン（１）とする。また、キャッシュメモリ７内に書き込む新データ（ＮＤ＃１）を保持する場合は、書き込み後の２つの論理アドレス４５内のキャッシュアドレス４７に、キャッシュメモリ７内の新データ（ＮＤ＃１）が格納されているアドレスを登録し、キャッシュフラグ４８をオン（１）とする。この新データ（ＮＤ＃１）をキャッシュメモリ７上に残さない場合、ＭＰ１　２０はこの報告を元にアドレステーブル４０のキャッシュフラグ４８をオフ（０）にする。さらに、書き込みを行った論理グループ１０内のＳＣＳＩ内Ａｄｄｒ４４に対し、スペースドライブ番号５１のＳＤフラグをオン（１）とする。

【００６２】上記のようにアドレステーブル４０の更新が完了し、ＭＰ１　２０はＳＤ＃４とＳＤ＃５の両方のＳＣＳＩドライブ１２からの完了報告を受け取った後、ＣＰＵ１に対し擬似的に書き込み処理の終了報告を行う。ＳＤ＃４，５のＳＣＳＩドライブ１２に対する新データ（ＮＤ＃１）の格納は完了しても、キャッシュメモリ７内には新データ（ＮＤ＃１）が存在しており、パリティの更新はキャッシュメモリ７内に格納されている新データ（ＮＤ＃１）で行なう。

【００６３】以上のようにＭＰ１　２０がＣＰＵ１に対し擬似的に終了報告を行った後は、ＣＰＵ１は書き込み処理を終了したと認識しているが、ＭＰ１　２０は新しいパリティを作成して、当該ＳＣＳＩドライブ１２書き込んでいないため、まだ書き込み処理は終了していない。そこで、ＭＰ１　２０がＣＰＵ１に対し終了報告を行った後に、ＭＰ１　２０が独自に新しいパリティを作成し、当該ＳＣＳＩドライブ１２に書き込む。この方法について以下に説明する。

【００６４】ＭＰ１　２０はＣＰＵ１に対し書き込み処理の終了報告を行った後、図５に示すようにＣＰＵ１からの読みだし、書き込み要求の状況（ＩＯ状況）を監視する（５１０）。具体的には、ＭＰ１　２０が当該論理グループ１０に対するＣＰＵ１からの読みだし、書き込み要求の単位時間当りの回数を数える。当該論理グループ１０に対するこの回数が予めユーザまたはシステム管理者により設定された数より少なく、パリティの作成およびこの作成したパリティを書き込むＳＣＳＩドライブ１２が所属する論理グループ１０に対し、読みだし、書き込み要求がＣＰＵ１より発行されていない場合、パリティの作成およびパリティの当該ＳＣＳＩドライブ１２への書き込み処理を開始する。

【００６５】新しくパリティを作成する場合、ＣＰＵ１から書き込み先に指定されたアドレスに書き込まれているデータ（旧データ）と、更新するパリティ（旧パリティ）を読みだし、新しく作成したパリティ（新パリティ）をＳＣＳＩドライブ１２に書き込む。この時、旧データと旧パリティを読みだすＳＣＳＩドライブ１２と新パリティを書き込む当該ＳＣＳＩドライブ１２に対し、ＭＰ１　２０はＣＰＵ１からの読みだし、書き込み要求と同様な擬似的な読みだし、書き込み要求を発行する。この、擬似的な読みだし、書き込み要求が発行されているＳＣＳＩドライブ１２に対しＣＰＵ１から読みだしまたは書き込み要求が発行された場合は、ＭＰ１　２０はＣＰＵ１からの読みだし、書き込み要求を受付、処理の待ち行列とする。

【００６６】次に新パリティの作成および当該ＳＣＳＩドライブ１２への作成したパリティの書き込み方法を具体的に示す。ＭＰ１　２０はＳＤ＃１のＳＣＳＩドライブ１２に対して旧データの読み出しと、ＳＤ＃６のＳＣＳＩドライブ＃１２に対して旧パリティの読み出し要求を発行するようにＤｒｉｖｅ　ＩＦ２８に指示する（５１４）。

【００６７】Ｄｒｉｖｅ　ＩＦ２８から読み出しコマンドを発行された当該ＳＣＳＩドライブ１２ではＤｒｉｖｅ　ＩＦ２８から送られてきたＳＣＳＩ内Ａｄｄｒ４４へシーク、回転待ちのアクセス処理を行なう。キャッシュメモリ７内には新データ（ＮＤ＃１）が存在しており、パリティの更新はキャッシュメモリ７内に格納されている新データ（ＮＤ＃１）で行なう。

【００６８】もし、キャッシュメモリ７に新データ（ＮＤ＃１）が存在していない場合はスペース領域に二重化して書き込まれているデータをキャッシュメモリ７に読みだす。

【００６９】ＳＤ＃１，６のＳＣＳＩドライブ１２においてシーク、回転待ちのアクセス処理が完了し旧データ（Ｄ＃１）および旧パリティ（Ｐ＃１）の読み出しが可能になり次第、旧データ（Ｄ＃１）および旧パリティ（Ｐ＃１）を読み出し、キャッシュメモリ７に格納する。ＳＤ＃１のＳＣＳＩドライブ１２から旧データ（Ｄ＃１）を、ＳＤ＃６のＳＣＳＩドライブ１２から旧パリティ（Ｐ＃１）を読み出し、それぞれをキャッシュメモリ７に格納した後、ＭＰ１　２０はキャッシュメモリ７内に格納されている書き込む新データ（ＮＤ＃１）とで排他的論理和により、更新後の新パリティ（ＮＰ＃１）

を作成するようにＰＧ３６に指示を出し、ＰＧ３６において新パリティ（ＮＰ＃１）を作成しキャッシュメモリ７に格納する（５１６）。

【００７０】新パリティ（ＮＰ＃１）をキャッシュメモリ７に格納した後、ＭＰ１　２０では新パリティ（ＮＰ＃１）を格納する論理アドレス４５のキャッシュアドレス４７にキャッシュメモリ７内の新パリティ（ＮＰ＃１）を格納したアドレスを登録し、キャッシュフラグ４８をオン（１）とし、無効フラグ４９とドライブフラグ５２をオフ（０）とする（５１８）。新パリティ（ＮＰ＃１）の作成が完了したことを認識し、ＳＤ＃６のＳＣＳＩドライブ１２にＩＯ要求が発行されていない時に、更新後の新パリティ（ＮＰ＃１）を書き込むようにＤｒｉｖｅ　ＩＦ２８に対し指示する。

【００７１】ＳＤ＃６のＳＣＳＩドライブ１２への更新後の新パリティ（ＮＰ＃１）の書込み方法（５２０）は、先に述べた書き込む新データ（ＮＤ＃１）をＳＤ＃４，５のＳＣＳＩドライブ１２に書き込んだ方法と同じである。新パリティ（ＮＰ＃１）の作成が完了したらＭＰ１　２０はＤｒｉｖｅ　ＩＦ２８にＳＤ＃６のＳＣＳＩドライブ１２に対し、書込みコマンドを発行するように指示し、当該ＳＣＳＩドライブ１２では指示ＳＣＳＩ内Ａｄｄｒ４４へシーク、回転待ちのアクセス処理を行う。新パリティ（ＮＰ＃１）は既に作成されキャッシュメモリ７に格納されており、ＳＤ＃６のＳＣＳＩドライブ１２におけるアクセス処理が完了した場合、Ｃ　Ａｄｐ１４はキャッシュメモリ７から新パリティ（ＮＰ＃１）を読み出してＤｒｉｖｅ　ＩＦ２８へ転送する。Ｄｒｉｖｅ　ＩＦ２８では転送されてきた新パリティ（ＮＰ＃１）をドライブユニットパス９－１から４の中の１本を介してＳＤ＃６のＳＣＳＩドライブ１２へ転送する（５２２）。

【００７２】新パリティ（ＮＰ＃１）のＳＤ＃６のＳＣＳＩドライブ１２への書込みが完了すると、ＳＤ＃６のＳＣＳＩドライブ１２はＤｒｉｖｅ　ＩＦ２８に完了報告を行ない、Ｄｒｉｖｅ　ＩＦ２８がこの完了報告を受け取ったことを、ＭＰ１　２０に報告する。この時、ＭＰ１　２０は、この新データ（ＮＤ＃１）をキャッシュメモリ７上に残さない場合は、この報告を元にアドレステーブル４０のキャッシュフラグ４８をオフ（０）にし、キャッシュメモリ７上に残す場合はオン（１）とする。さらに、アドレステーブル４０において、書き込み後の新パリティ（ＮＰ＃１）の論理アドレス４５の無効フラグをオフ（０）とし、ドライブフラグ５２をオン（１）とする（５２４）。

【００７３】このような、新しく作成した新パリティ（ＮＰ＃１）を当該ＳＣＳＩドライブ１２に書き込んだ後は、旧データ（Ｄ＃１）が格納されていたＳＤ＃１のＳＣＳＩドライブ１２の旧データ（Ｄ＃１）と二重化されている新データ（ＮＤ＃１）が格納されているＳＤ＃４，５のＳＣＳＩドライブ１２の中でＳＣＳＩドライブ番号の小さいＳＤ＃４のＳＣＳＩドライブ１２に格納されている新データ（ＮＤ＃１）をスペース領域として解放し、次の書き込み処理用にスペース領域として登録する。この登録の方法は、ＭＰ１　２０がアドレステーブル４０においてＳＤ＃１とＳＤ＃４のＳＣＳＩ内Ａｄｄｒ４４がＤ　ＡＤＲ１の旧データ（Ｄ＃１）と、二重化されていた新データ（ＮＤ＃１）の一方が格納されている論理アドレス４５において、無効フラグをオン（１）とし、さらにスペースドライブ番号５１にＳＤ＃１とＳＤ＃４を登録し、ＳＤフラグをオフ（０）にする（５２６）。

【００７４】以上のように書き込み処理時に書き込む新データ（ＮＤ＃１）を一旦二重化して論理グループ１０内に格納しておき、後で比較的ＣＰＵ１からの読みだし、書き込み要求が少ないときに、新パリティ（ＮＰ＃１）を作成し、ＳＣＳＩドライブ１２に格納することで、書き込み処理時の応答時間が短縮され、従来のように新パリティ（ＮＰ＃１）の作成により他の読みだし、書き込み処理がまたされることが減少する。

【００７５】もし、新パリティ（ＮＰ＃１）を当該ＳＣＳＩドライブ１２への書き込み前に論理グループ１０内のＳＣＳＩドライブ１２に障害が発生した場合は図１４に示すように回復処理を行う。図４（ａ）に示すようにＳＤ＃６のＳＣＳＩドライブ１２に対し新パリティ（ＮＰ＃１）を書き込む前に（１４０２）、ＳＤ＃１，２，３の内のどれか１台のＳＣＳＩドライブ１２に障害が発生した場合（１４０６）、障害が発生していないＳＣＳＩドライブ１２からのデータと、旧パリティから障害が発生したＳＣＳＩドライブ１２内のデータは回復することが可能となる（１４１０）。例えばＳＤ＃１のＳＣＳＩドライブ１２に障害が発生した場合、ＳＤ＃２，３のＤ＃２，Ｄ＃３とＳＤ＃６のＳＣＳＩドライブ１からの旧パリティ（Ｐ＃１）からＳＤ＃１のＳＣＳＩドライブ１２内のデータであるＤ＃１を回復することが可能となる。また、新データ（ＮＤ＃１）が二重化して格納されているＳＤ＃４，５のどちらか一方に障害が発生した場合は、二重化データの一方のデータにより回復することが可能となる（１４１２）。

【００７６】図４（ｂ）に示すようにＳＤ＃６のＳＣＳＩドライブ１２に対し新パリティ（ＮＰ＃１）を書き込んだ後はＳＤ＃２，３，５のＳＣＳＩドライブ１２内のデータ（Ｄ＃２，３，ＮＤ＃１）に対し新パリティ（ＮＰ＃１）が作成されＳＤ＃６のＳＣＳＩドライブ１２に格納されており、ＳＤ＃２，３，５のＳＣＳＩドライブ１２の中の１台のＳＣＳＩドライブ１２に障害が発生した場合、残りの正常なＳＣＳＩドライブ１２内のデータとＳＤ＃６のＳＣＳＩドライブ１２内のパリティにより、障害が発生したＳＣＳＩドライブ１２内のデータは回復される。

【００７７】例えばＳＤ＃２のＳＣＳＩドライブ１２に障害が発生した場合、ＳＤ＃２のＳＣＳＩドライブ１２内のデータ（Ｄ＃２）は、ＳＤ＃３のＳＣＳＩドライブ１２内のデータ（Ｄ＃３）とＳＤ＃４のＳＣＳＩドライブ１２内のデータ（ＮＤ＃１）と、ＳＤ＃６のＳＣＳＩドライブ１２内のパリティ（ＮＰ＃１）から障害が発生したＳＤ＃２のＳＣＳＩドライブ１２内のデータ（Ｄ＃２）は回復される。

【００７８】本発明のように書き込み処理においてスペース領域に書き込みデータ（新データ）をとりあえず二重化して格納し、この段階でＣＰＵ１に対し書き込み処理の終了報告を行うことにより、ＣＰＵ１にとってはこの二重化してＳＣＳＩドライブ１２に書き込む時間が書き込み処理時間になる。従来のアレイディスクでは図１２に示すように書き込み時に平均１．５回転の回転待ち時間が必要としたのが、もし、論理グループ１０を構成するＳＣＳＩドライブ１２の回転を同期させた場合は、回転待ちは平均０．５回転となる。また、新パリティをＳＣＳＩドライブ１２に書き込む前に論理グループ１０を構成するＳＣＳＩドライブ１２に障害が発生しても、先に述べたように旧パリティと二重化された新データにより従来のアレイディスクと同様に、障害回復を行うことが可能となる。

【００７９】本実施例ではパリティおよびパリティの作成に関与したデータとスペース領域は、論理グループ１０を構成する各ＳＣＳＩドライブ１２において、各ＳＣＳＩ内Ａｄｄｒ４４を同一とした。しかし、アドレステーブル４０の各論理アドレス４５とパリティドライブ番号５０、スペースドライブ番号５１において、論理グループ１０のアドレスを付加することにより、論理グループ１０を構成する各ＳＣＳＩドライブ１２においてＳＣＳＩ内Ａｄｄｒ４４を任意とすることが可能となる。

【００８０】本実施例では書き込み時の回転待ち時間を減少させる目的で、書き込みデータを一旦二重化してＳＣＳＩドライブ１２に書き込み、後の適当なタイミングでパリティを更新し、パリティの更新後は二重化した書き込みデータの一方を開放する方法について説明した。本発明は、上記のような性能向上の観点とは別に、以下に示すような応用例も考えられる。

【００８１】パリティによる信頼性と、二重化による信頼性では、二重化による信頼性の方が信頼性向上のために必要とする容量は大きいが、信頼性は高くなる。この特徴を活かし、本発明の応用として、最新に書き込まれたデータまたは、頻繁に書き込み要求が発行されるデータについては、二重化されているため、高信頼とし、あまり書き込み要求が発行されないデータについては、二重化ほど高信頼ではないが、信頼性向上のために必要とする容量が二重化と比較して少なくてすむ、パリティにより信頼性を確保する。つまり、あまり書き込み要求が発行されないデータに対しては、信頼性は二重化ほど高くないが、パリティを付けるだけですむ、パリティで信頼性を確保し、最新に書き込まれたデータまたは、頻繁に書き込み要求が発行されるデータについては、パリティと比較し信頼性向上のために容量を多く必要とするが、高信頼である二重化で高信頼とする、二段階の信頼性のレベルを設定することも可能である。

【００８２】〈実施例２〉本発明の他の実施例を図６を中心にして説明する。本実施例では実施例１で示したシステムにおいて、ＳＣＳＩドライブ１２に障害が発生した時に、その障害が発生したＳＣＳＩドライブ１２内のデータを回復し、それを格納するための領域にスペース領域を使用する例を示す。

【００８３】本発明では図３に示すように、論理グループ１０内の各ＳＣＳＩドライブ１２では、その内部の各々対応する同一ＳＣＳＩ内Ａｄｄｒ４４のデータによりパリティグループを構成する。具体的にはＳＤ＃１，２，３のＳＣＳＩドライブ１２内のＤａｔａ＃１、２、３（Ｄ＃１，２，３）でＰＧ３６によりパリティ＃１（Ｐ＃１）が作られＳＤ＃６のＳＣＳＩドライブ１２内に格納される。本実施例ではパリティは奇数パリティとし、Ｄａｔａ＃１，２，３（Ｄ＃１，２，３）の各データにおける各々対応するビットについて１の数を数え、奇数であれば０、偶数であれば１とする（排他的論理和）。もし、ＳＤ＃１のＳＣＳＩドライブ１２に障害が発生したとする。この時、Ｄａｔａ＃１（Ｄ＃１）は読み出せなくなる。

【００８４】本実施例ではパリティグループ当りにパリティを１個しか持っていないため、１台のＳＣＳＩドライブ１２の障害はデータを復元できるが、データの復元が完了する前に更にもう一台のＳＣＳＩドライブ１２に障害が発生した場合復元出来ない。そこで、この様な場合、２台目のＳＣＳＩドライブ１２障害が発生する前に、残りのＤａｔａ＃２、３（Ｄ＃２，３）とパリティ＃１（Ｐ＃１）をキャッシュメモリ７に転送し、ＭＰ１　２０はＰＧ３６に対しＤａｔａ＃１（Ｄ＃１）を復元する回復処理を早急に行なうように指示する。この回復処理を行ないＤａｔａ＃１（Ｄ＃１）を復元した後は、ＭＰ１　２０はこのＤａｔａ＃１（Ｄ＃１）をＳＤ＃４または６のどちらかのスペース領域に格納する。

【００８５】これにより、スペース領域を実施例１で示したような、書き込み処理時の回転待ち時間を短縮させるためだけではなく、ＳＣＳＩドライブ１２に障害が発生したときに、復元したデータを格納するためのスペアー領域としても活用する。この様に、ＭＰ１　２０がスペース領域に回復したＤａｔａ＃１（Ｄ＃１）を格納した後は、キャッシュメモリ７にある図３に示すアドレステーブル４０において、スペースドライブ番号５１の中で、回復データを格納した方のスペースドライブ番号５１を削除し、この削除したドライブ番号に対する論理アドレス４５に、回復したＤａｔａ＃１（Ｄ＃１）の論理

アドレス４５の内容を複写する。

【００８６】図６に示すようにＳＤ＃１のＳＣＳＩドライブ１２にはＤａｔａ＃１（Ｄ＃１）の他にスペース領域、パリティ、Ｄａｔａ＃１３、１６、１９、２２（Ｄ＃１３，１６，１９，２２）が格納されている。スペース領域については回復処理を行ない復元する必要は無い。パリティ＃３（Ｐ＃３）はＳＤ＃３，４，５のＳＣＳＩドライブ１２からＤａｔａ＃７、８、９（Ｄ＃７，８，９）を読み出して新たに作成しＳＤ＃２か６のＳＣＳＩドライブ１２のスペース領域に格納する。Ｄａｔａ＃１３（Ｄ＃１３）はＳＤ＃３，５，６のＳＣＳＩドライブ１２からパリティ＃５（Ｐ＃５）、Ｄａｔａ＃１４、１５（Ｄ＃１４，１５）を読み出して、回復処理を行ない復元し、ＳＤ＃２または４のＳＣＳＩドライブ１２のスペース領域に格納する。Ｄａｔａ＃１６（Ｄ＃１６）はＳＤ＃２，４，６のＳＣＳＩドライブ１２からＤａｔａ＃１７（Ｄ＃１７）、パリティ＃６（Ｐ＃６）、Ｄａｔａ＃１８（Ｄ＃１８）を読み出して、回復処理を行ない復元し、ＳＤ＃３または５のＳＣＳＩドライブ１２のスペース領域に格納する。以下同様にＤａｔａ＃１９、２２（Ｄ＃１９，２２）と回復処理を行ない論理グループ１０内のスペース領域に格納していく。

【００８７】この様に、ＳＤ＃２，３，４，５，６のＳＣＳＩドライブ１２内のスペース領域に、ＳＤ＃１のＳＣＳＩドライブ１２の回復データを全て格納した後は、スペース領域が論理グループ１０において一つしかないため、実施例１で述べたような書き込み時の回転待ちを短くすることは出来ないため、従来のアレイディスクであるＲＡＩＤのレベル５の処理となる。また、ＳＤ＃１のＳＣＳＩドライブ１２の回復データを全て格納した後は、ＳＤ＃２，３，４，５，６のＳＣＳＩドライブ１２の中で更にもう一台ＳＣＳＩドライブ１２に障害が発生した場合、同様にその障害が発生したＳＣＳＩドライブ１２内のデータについて回復処理を行ない、残りのスペース領域に格納し、処理を行なえる。

【００８８】この様にして、論理グループ１０内のスペース領域を全て使いきってしまった場合は、障害ＳＣＳＩドライブ１２を正常のＳＣＳＩドライブ１２に交換し、この交換した正常なＳＣＳＩドライブ１２は全てスペース領域として論理グループを再構成する。

【００８９】障害ＳＣＳＩドライブ１２を正常のＳＣＳＩドライブ１２に交換した直後は、スペース領域が特定ＳＣＳＩドライブ１２に集中した形になっているため、このＳＣＳＩドライブ１２が使用出来ずにまたされることが多くなりネックとなるため、実施例１で示した回転待ち時間を短縮する効果が、効率的に発揮出来ない。しかし、時間が立つにつれて、スペース領域が分散されてＳＣＳＩドライブ１２障害前の状態に戻っていき、次第に解消されていく。もし、この時間が問題となる場合は、ＳＣＳＩドライブ１２に障害が発生したことを感知した場合、正常なＳＣＳＩドライブ１２に交換して、この交換した正常なＳＣＳＩドライブ１２に障害が発生したＳＣＳＩドライブ１２内のデータとパリティをユーザが復元することも可能とする。なお、この時スペース領域に関しては復元せずにスペース領域として空けておく。

【００９０】本実施例ではこの回復処理と、スペース領域へ復元したデータを書き込む処理をＭＰ１　２０が独自に行なう。この様に独自に行なうことによりＳＣＳＩドライブ１２に障害が発生した場合、障害が発生したＳＣＳＩドライブ１２を正常なＳＣＳＩドライブ１２に交換し回復したデータを書き込むのと比較し、本発明ではシステムを使用するユーザがＳＣＳＩドライブ１２に障害が発生するとすぐに正常なＳＣＳＩドライブ１２と交換する必要が無いため、ユーザの負担が軽くなる。

【００９１】〈実施例３〉本発明の第三の実施例を図７～図１１により説明する。本実施例では図７、８に示すように論理グループ１０単位にサブＤＫＣ１１を設け、その内部に図９に示すように実施例１、２において示したキャッシュメモリ７内のアドレステーブル４０とＲＰＣ２７、ＰＧ３６、サブキャッシュ３２とそれらを制御するマイクロプロセッサＭＰ３　２９を持たせたものを示す。本実施例におけるデータの処理手順は実施例１および２で示したものと同様である。以下には実施例１、２で示した処理手順と異なる部分のみを図１０および図１１を用いて示す。本実施例では図９に示すように実施例１、２で示したキャッシュメモリ７内のアドレステーブル４０をサブＤＫＣ１１内のデータアドレステーブル（ＤＡＴ）３０に格納する。ＤＡＴ３０は格納されているテーブルの形式や機能は実施例１、２と同様であるが、異なるのはデータを格納するＳＣＳＩドライブアドレス４２が論理グループ１０に限られている点と、メモリがアドレステーブル４０を格納するデータを格納するのとは別の専用メモリである。ＡＤＣ２内のＧＡＴ２３はＣＰＵ１から指示されたＣＰＵ指定アドレスから、そのＣＰＵ指定アドレスが指示する場所がＡＤＵ３内のどの論理グループ１０かを判定するのみである。キャッシュメモリ７内には、その特別な領域に図１０に示すような論理グループテーブル（ＬＧＴ）６０が格納されている。

【００９２】ＬＧＴ６０は図１０に示すようにＣＰＵ１から指定されるＣＰＵ指定ドライブ番号４１とＣＣＨＨＲ４６に対応して、論理グループアドレス６１が決定できるテーブルとなっている。また、ＬＧＴ６０にはＣＰＵ指定アドレスに対応するデータがキャッシュメモリ７内に存在する場合、そのデータのキャッシュメモリ７内のアドレスをキャッシュアドレス４７に登録でき、また、キャッシュメモリ７内にデータが存在する場合オン（１）とし、キャッシュメモリ７内に存在しない場合オフ（０）とするキャッシュフラグ４８が用意されてい

る。ユーザは初期設定する際に自分の使用可能な容量に対する領域を確保するが、その際にＡＤＣ２のＭＰ１２０が、ＬＧＴ６０により論理グループ１０を割当てる。この時、ＭＰ１　２０はＬＧＴ６０にユーザが確保するために指定したＣＰＵ指定アドレスに対応する領域を登録する。

【００９３】そこで、実際の読みだし、書き込み処理においては、ＧＡＴ２３はＬＧＴ６０によりＣＰＵ１から指定してきたＣＰＵ指定アドレスに対応した論理グループ１０を認識することが可能となる。読み出し時はＧＡＴ２３がＬＧＣにより論理グループ１０を確定し、その確定した結果をＭＰ１　２０に報告し、ＭＰ１　２０はこの当該論理グループ１０に対し読み出し要求を発行するようにＤｒｉｖｅＩＦ２８に指示する。ＭＰ１　２０から指示を受けたＤｒｉｖｅ　ＩＦ２８は当該論理グループ１０のサブＤＫＣ１１に対し読み出し要求とＣＰＵ１が指定するＣＰＵ指定アドレスを発行する。サブＤＫＣ１１ではマイクロプロセッサであるＭＰ３　２９がこの読み出し要求のコマンドとＣＰＵ指定アドレスを受け付け、実施例１と同様に、Ｄｒｉｖｅ　ＩＦ２８から送られてきたＣＰＵ指定アドレスをＤＡＴ３０を参照し、当該データが格納されている論理グループ１０内の論理アドレス４５に変換し、この論理アドレス４５から当該ＳＣＳＩドライブアドレス４２（ＳＣＳＩドライブ番号４３とその中のＳＣＳＩ内Ａｄｄｒ４４）を確定する。

【００９４】このアドレスの確定後ＭＰ３　２９は当該ＳＣＳＩドライブ１２に対し、読み出し要求を発行する。ＭＰ３　２９から読み出し要求を発行されたＳＣＳＩドライブ１２ではＳＣＳＩ内Ａｄｄｒ４４へシーク、回転待ちを行ない、当該データの読み出しが可能になり次第、当該データをドライブアダプタ回路（Ｄｒｉｖｅ Ａｄｐ）３４に転送し、Ｄｒｉｖｅ　Ａｄｐ３４はサブキャッシュメモリ３２に格納する。サブキャッシュメモリ３２に当該データの格納が完了した後、Ｄｒｉｖｅ Ａｄｐ３４はＭＰ３　２９に格納報告を行ない、ＭＰ３　２９はＤＡＴ３０の当該データの当該論理アドレス４５内の当該キャッシュフラグ４８をオン（１）とする。後に当該キャッシュフラグ４８がオン（１）のデータに対し読み出しまたは書込み要求が発行された場合は、サブキャッシュ３２内で処理を行なう。実施例１と同様にＭＰ３　２９によるＤＡＴ３０の更新が終了すると、ＭＰ３　２９はＡＤＣ２内のＤｒｉｖｅ　ＩＦ２８に対しデータ転送可能という応答を行ない、Ｄｒｉｖｅ　ＩＦ２８はこの応答を受け取ると、ＭＰ１　２０に対し報告する。

【００９５】ＭＰ１　２０はこの報告を受け取ると、キャッシュメモリ７への格納が可能なら、Ｄｒｉｖｅ　ＩＦ２８に対しサブＤＫＣ１１からデータを転送するように指示する。Ｄｒｉｖｅ　ＩＦ２８ではＭＰ１　２０からの指示を受けるとサブＤＫＣ１１のＭＰ３　２９に対し読み出し要求を発行する。この読み出し要求を受けたＭＰ３　２９はサブキャッシュアダプタ回路（ＳＣＡ）３１に対しサブキャッシュ３２から当該データを読みだすように指示し、ＳＣＡ３１は実際にデータを読み出してＤｒｉｖｅ　ＩＦ２８にデータを転送する。Ｄｒｉｖｅ　ＩＦ　２８がデータを受け取った後は、実施例１、２で示した処理を行なう。

【００９６】一方書込み時は読み出し時と同様に当該論理グループ１０を確定し、ＭＰ１２０はＤｒｉｖｅ　ＩＦ２８に対し当該論理グループ１０のＭＰ３　２９に対し書込み要求を発行するように指示する。当該論理グループ１０内のＭＰ３　２９は、書込み要求を受け付け、書込みデータをサブキャッシュ３２に格納した後は、図５のフローチャートに従って、実施例１、２と同様に処理を行う。本実施例では実施例１、２の効果を実現することが可能である。

【００９７】以上、磁気ディスク装置を用いたシステムを実施例として説明したが、本発明は光ディスク装置を用いたシステムにおいても同様な効果を発揮することが可能である。

【００９８】

【発明の効果】本発明によれば、データの書き込み時におけるパリティの更新処理をＣＰＵからの読みだしまたは書き込み要求が少ない時まで遅らせることが可能となる。これにより、ＣＰＵにとっては読みだしまたは書き込み処理要求が多い時は書き込み処理を高速に行え、これにより単位時間当りのＩ／Ｏ処理件数を増加させることが可能となる。さらに、通常は使用しないスペアのＳＣＳＩドライブを、回転待ち時間の短縮という、性能向上のために使用でき、ＳＣＳＩドライブ資源の有効活用が図れる。

【図面の簡単な説明】

【図１】第１の実施例の全体構成図。

【図２】第１の実施例のクラスタ内構成図。

【図３】アドレス変換テーブルの説明図。

【図４】書き込み処理時のデータ移動説明図。

【図５】書き込み処理フローチャート（１）。

【図６】データ回復処理説明図、パリティグループを構成する各データのディスク上での位置説明図。

【図７】第３の実施例の全体構成図。

【図８】第３の実施例のクラスタ内構成図。

【図９】第３の実施例のサブＤＫＣ内構成図。

【図１０】論理グループテーブル説明図。

【図１１】ＲＡＩＤ　Ｌｅｖｅｌ５における更新処理説明図。

【図１２】ＲＡＩＤ５における書き込み処理タイミングチャート。

【図１３】書き込み処理フローチャート（２）。

【図１４】パリティ作成処理フローチャート。

【符号の説明】

29

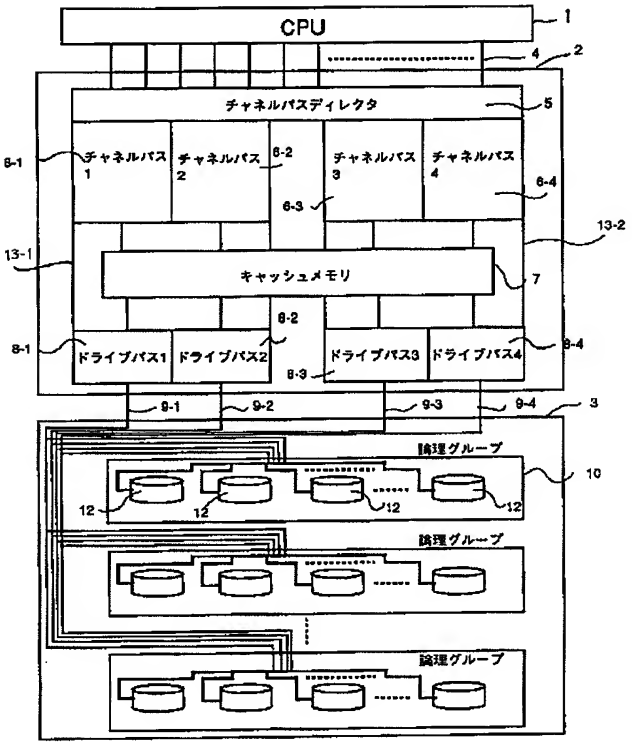
１…ＣＰＵ、２…アレイディスクコントローラ（ＡＤＣ）、３…アレイディスクユニット（ＡＤＵ）、４…外部インターフェースパス、５…チャネルパスディレクタ、６…チャネルパス、７…キャッシュメモリ、８…ドライブパス、９…アレイディスクユニットパス、１０…論理グループ、１２…ＳＣＳＩドライブ、１３…クラスタ、１４…ドライブ側キャッシュアダプタ（Ｃ　Ａｄｐ）、１５…インターフェースアダプタ、１６…チャネルパススイッチ、１７…制御信号線、１８…データ線、１９…パス、２０…マイクロプロセッサ１（ＭＰ１）、２１…チャネルインターフェース（ＣＨ　ＩＦ）回路、２２…データ制御回路（ＤＣＣ）、２３…グループアドレス変換回路（ＧＡＴ）、２４…チャネル側キャッシュアダプタ（Ｃ　Ａｄｐ）、２８…ドライブインターフェ＊

30

＊ース回路（Ｄｒｉｖｅ　ＩＦ）、２９…マイクロプロセッサ３（ＭＰ３）、３０…データアドレステーブル４０、３１…サブキャッシュアダプタ、３４…ドライブアダプタ（Ｄｒｉｖｅ　Ａｄｐ）、３５…ドライブパス、３６…パリティ生成回路、４０…アドレス変換用テーブル（アドレステーブル）、４１…ＣＰＵ指定ドライブ番号、４２…ＳＣＳＩドライブアドレス、４３…ＳＣＳＩドライブ番号、４４…ＳＣＳＩ内Ａｄｄｒ、４５…論理アドレス、４６…ＣＣＨＨＲ、４７…キャッシュアドレス、４８…キャッシュフラグ、４９…無効フラグ、５０…パリティドライブ番号、５１…スペースドライブ番号、５２…ドライブフラグ、５３…ＳＤフラグ、６０…論理グループテーブル、６１…論理グループアドレス

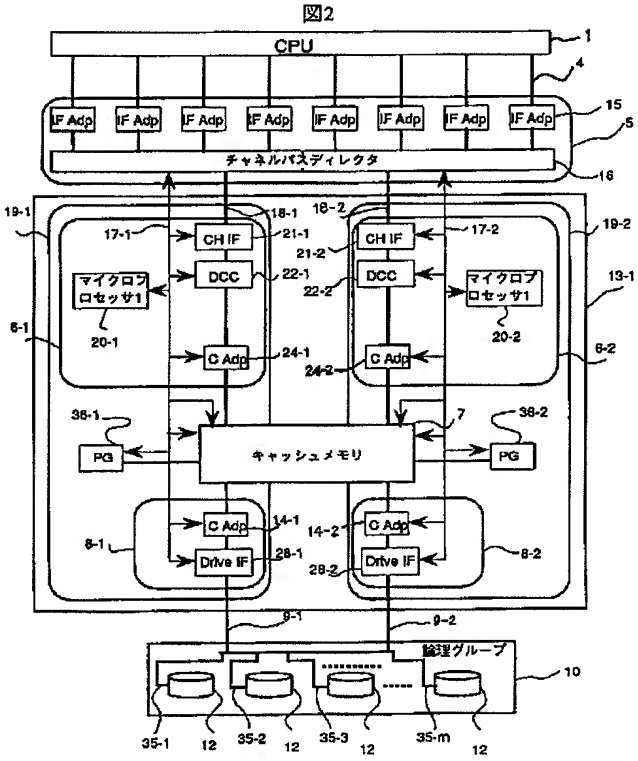
【図１】

図1



【図２】

図2



【図３】

図3

D: データ　P:パリティ　S:スペース領域

	SD#1	SD#2	SD#3	SD#4	SD#5	SD#6
D ADR1	D#1	D#2	D#3	S	S	P#1
D ADR2	S	D#4	D#5	D#6	S	P#2
D ADR3	P#3	D	D#7	D#8	D#9	S
D ADR4	S	P#4	S	D#10	D#11	D#12
D ADR5	D#13	S	P#5	S	D#14	D#15
D ADR6	D#16	D#17	S	P#6	S	D#18
D ADR7	D#19	D#20	S	S	D#21	P#7
D ADR8	D#22	D#23	D#24	P#8	S	S

CPU指定 Drive No.	SCSI 内Addr	PDrive No.	SDrive No.		SCSI ドライブアドレス / SCSI ドライブ番号 SD#1	SD#2	SD#3	SD#4	SD#5	SD#6
Drive#1	D ADR1	SD#6	SD#4	SD#5	Data#1	Data#2	Data#3	Space	Space	Parity#1
	D ADR2	SD#6	SD#1	SD#5	Space	Data#4	Data#5	Data#6	Space	Parity
	D ADR3	SD#1	SD#2	SD#6	Parity	Space	Data#7	Data#8	Data#9	Space
	D ADR4	SD#2	SD#1	SD#3	Space	Parity	Space	Data#10	Data#11	Data#12
	D ADR5	SD#3	SD#2	SD#4	Data#13	Space	Parity	Space	Data#14	Data#15
	D ADR6	SD#4	SD#3	SD#5	Data#16	Data#17	Space	Parity	Space	Data#18
	D ADR7	SD#6	SD#3	SD#4	Data#19	Data#20	Space	Space	Data#21	Parity
	D ADR8	SD#4	SD#	SD#6	Data#22	Data#23	#24	Parity	Space	Space

41 44 50 51 42 43 40

SCSI Drive No	SD Flag
SD#6	1

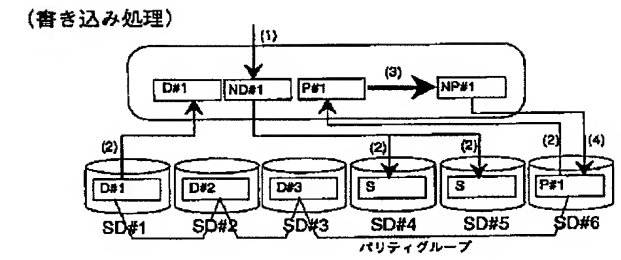
CCHHR	キャッシュアドレス	Cache Flag	無効 Flag	Drive Flag
ADR 8	C ADR2,1	1	0	1

43 45 52 53 46 47 48 49

PDrive No.:パリティドライブ番号
SDrive No.:スペースドライブ番号

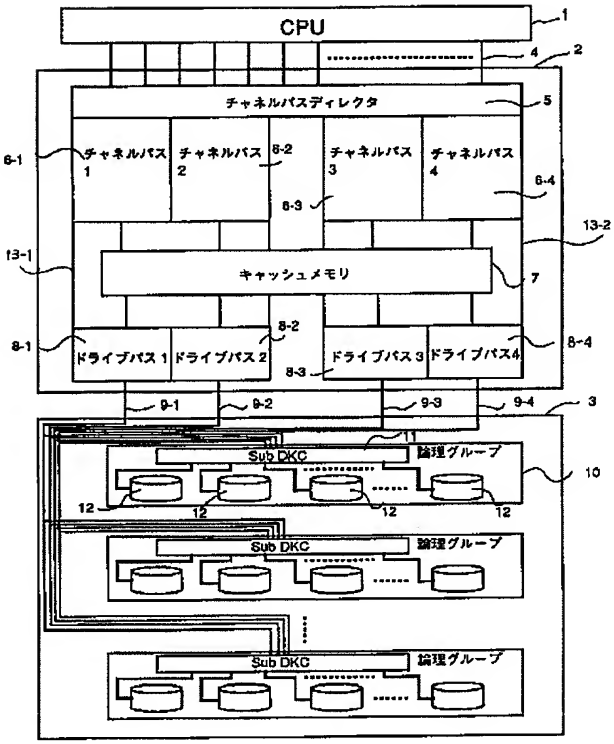
【図４】

図4



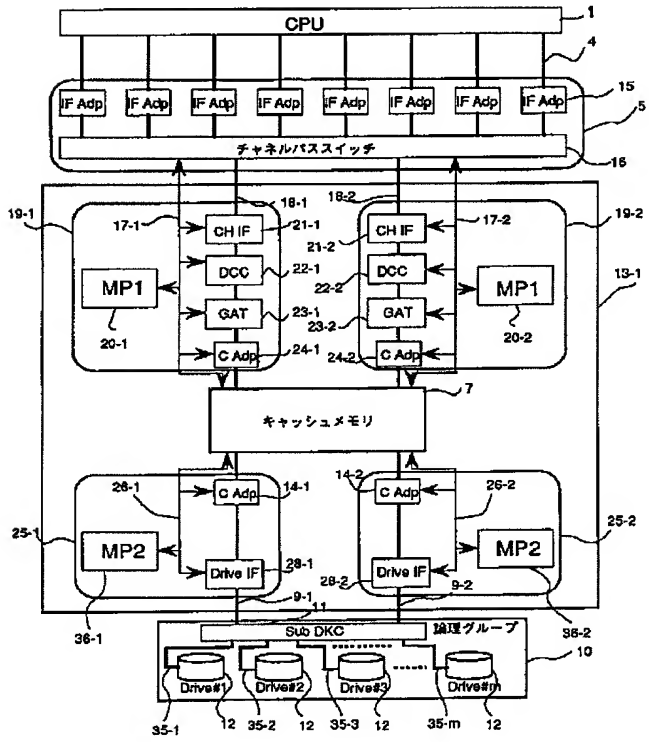
【図７】

図7



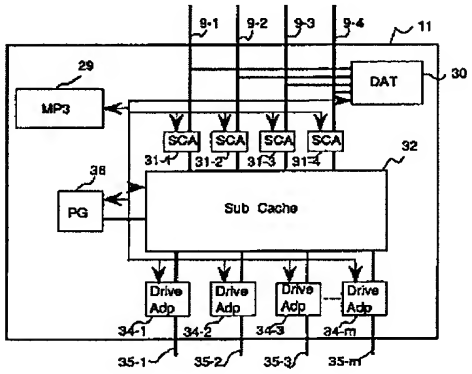
【図８】

図8



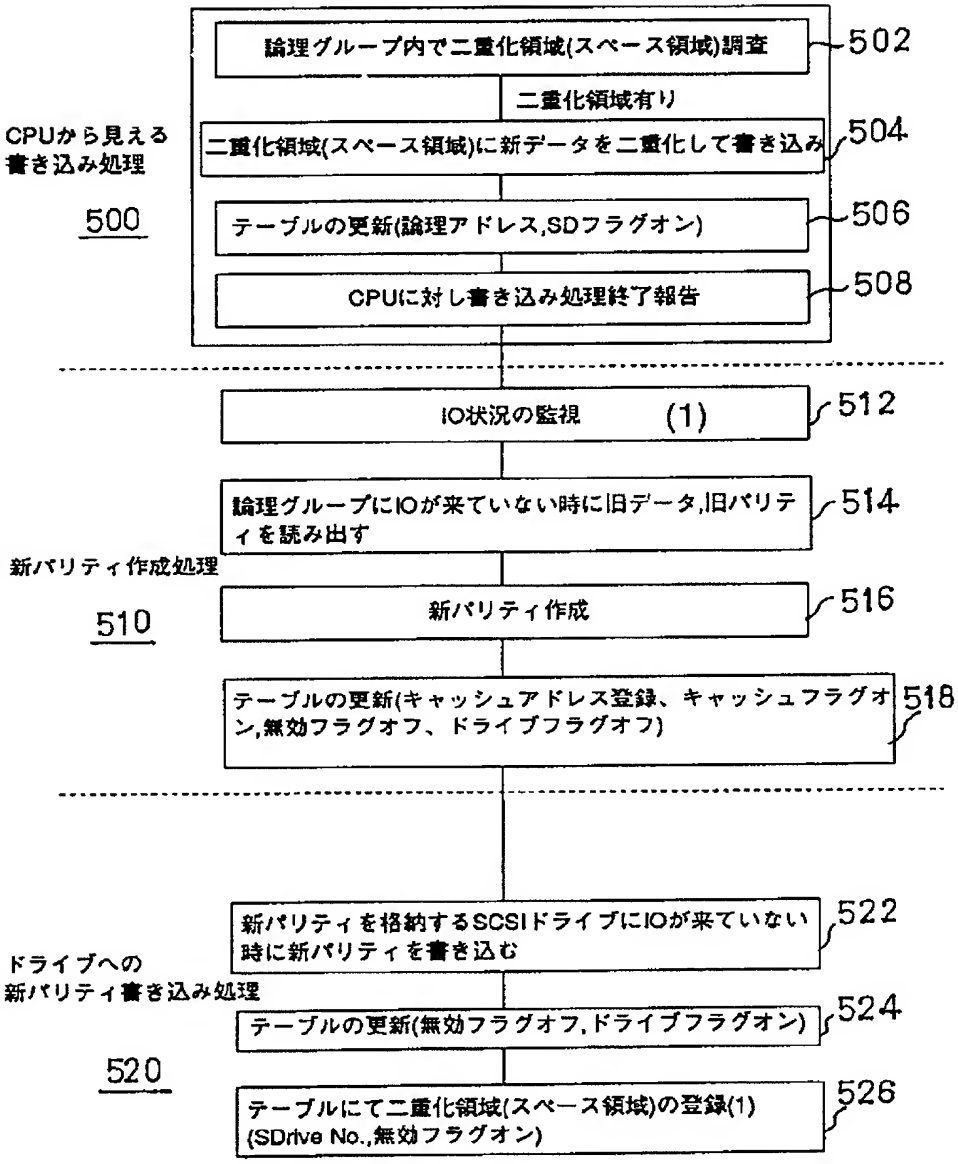
【図９】

図9



【図５】

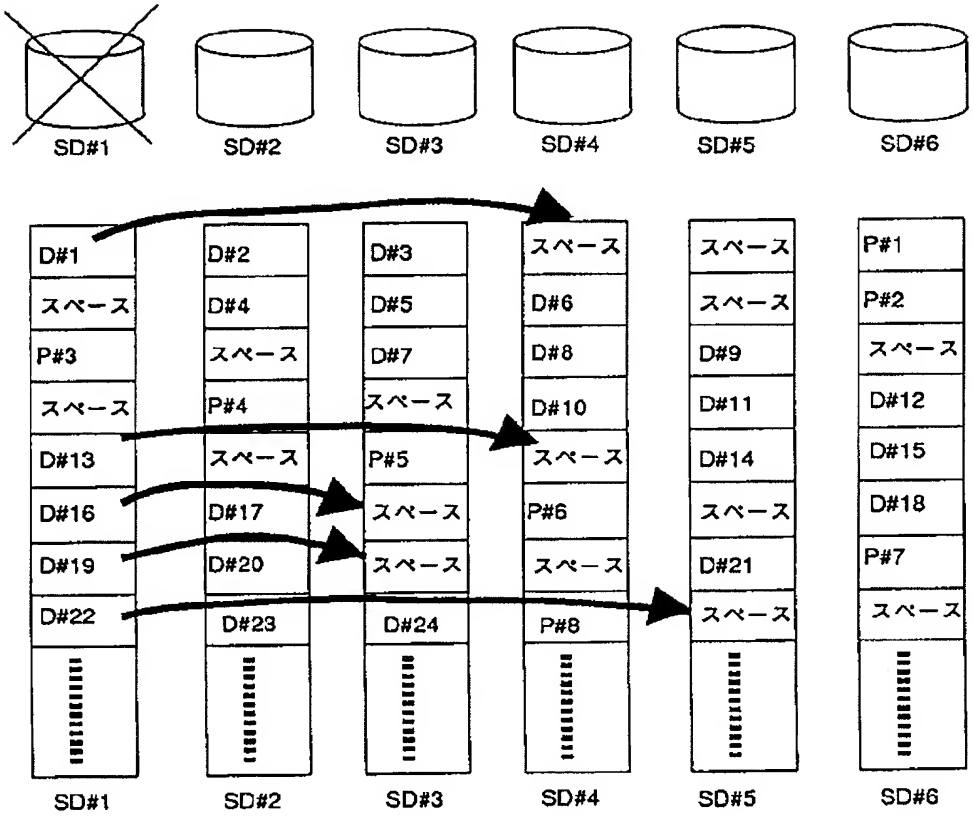
図5



(1)二重化領域は書き込み先に指定したアドレスに書き込まれているデータと二重化されているデータのペアの中でＳＣＳＩドライブ番号の小さい方とする

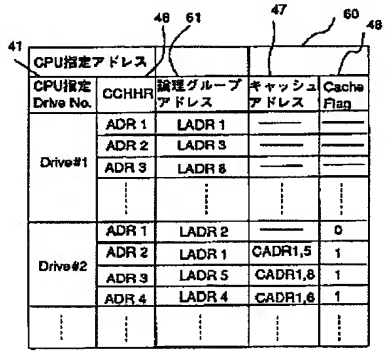
【図６】

図6


SD#1
SD#2
SD#3
SD#4
SD#5
SD#6
D#1
スペース
P#3
スペース
D#13
D#16
D#19
D#22
D#2
D#4
スペース
P#4
スペース
D#17
D#20
D#23
D#3
D#5
D#7
スペース
P#5
スペース
スペース
D#24
スペース
D#6
D#8
D#10
スペース
P#6
スペース
P#8
スペース
スペース
D#9
D#11
D#14
スペース
D#21
スペース
P#1
P#2
スペース
D#12
D#15
D#18
P#7
スペース
SD#1
SD#2
SD#3
SD#4
SD#5
SD#6

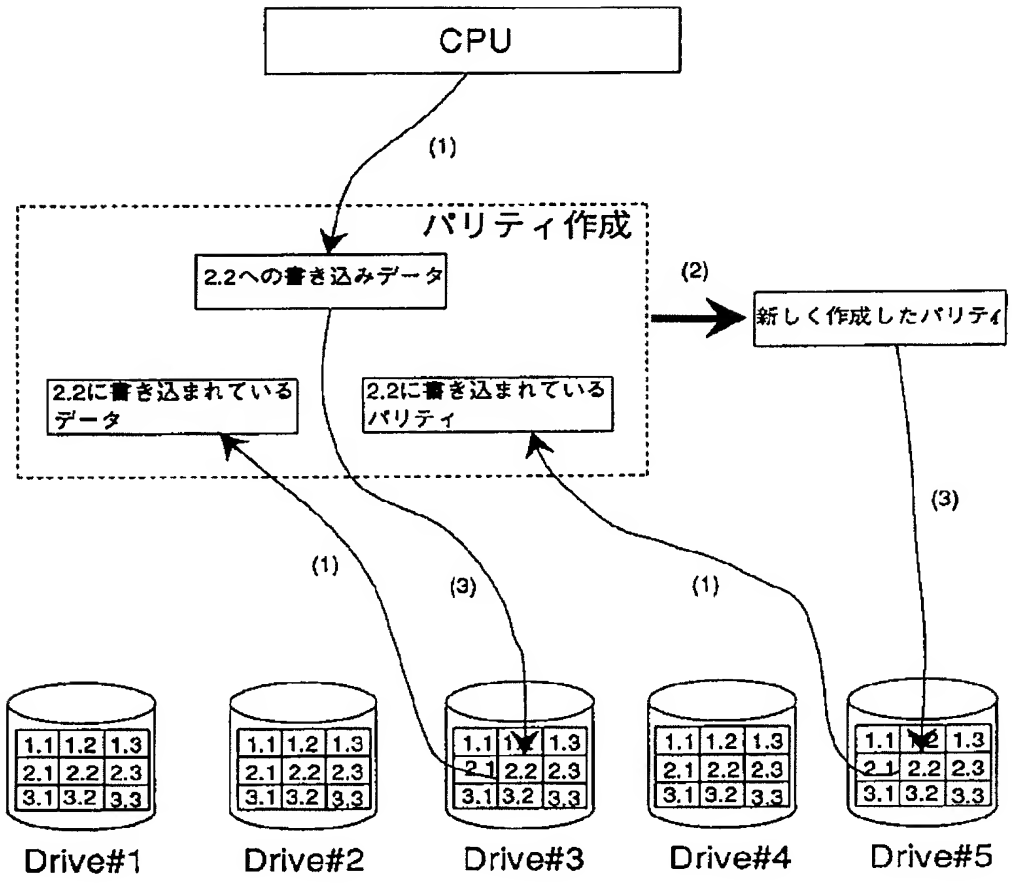
【図１０】

図10


41
48
61
47
60
48
CPU指定アドレス
CPU指定 Drive No.
CCHHR
論理グループアドレス
キャッシュアドレス
Cache Flag
Drive#1
ADR 1
ADR 2
ADR 3
LADR 1
LADR 3
LADR 6
Drive#2
ADR 1
ADR 2
ADR 3
ADR 4
LADR 2
LADR 1
LADR 5
LADR 4
CADR1,5
CADR1,6
CADR1,6
0
1
1
1

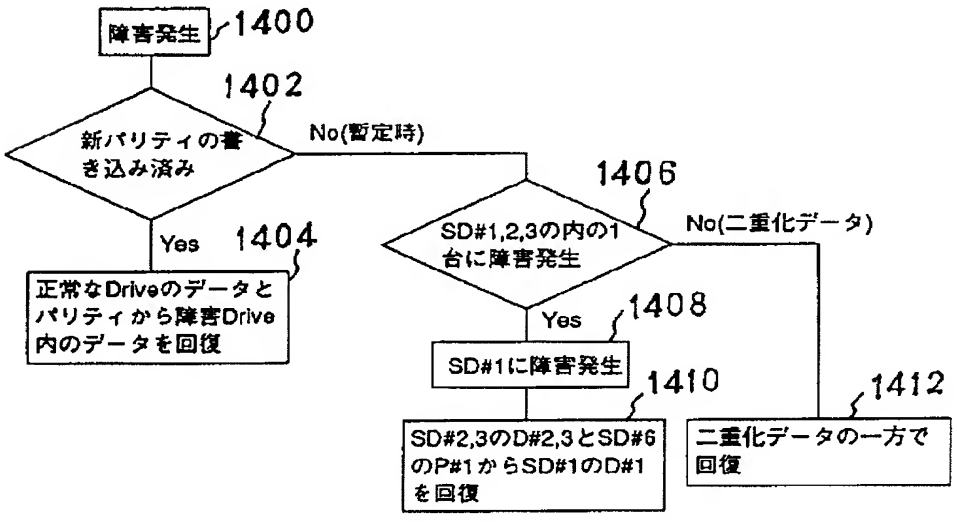
【図１１】

図11



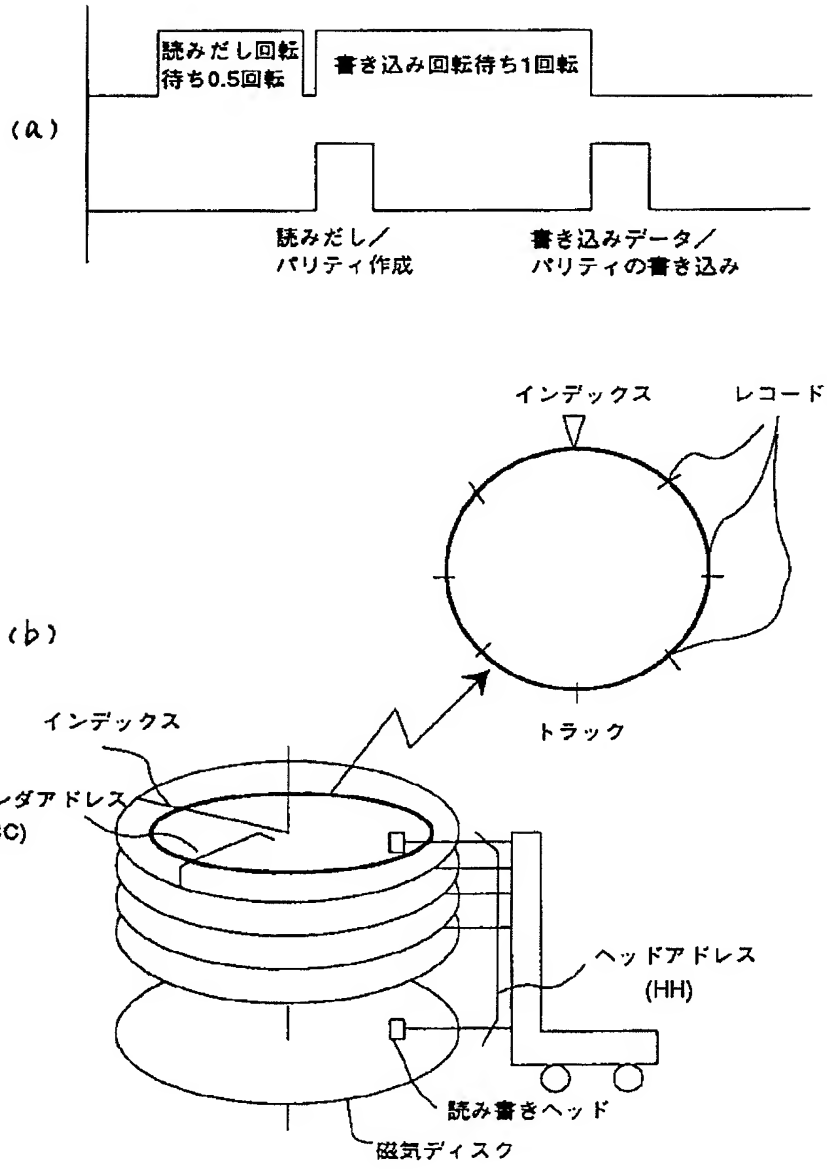
【図１４】

図14



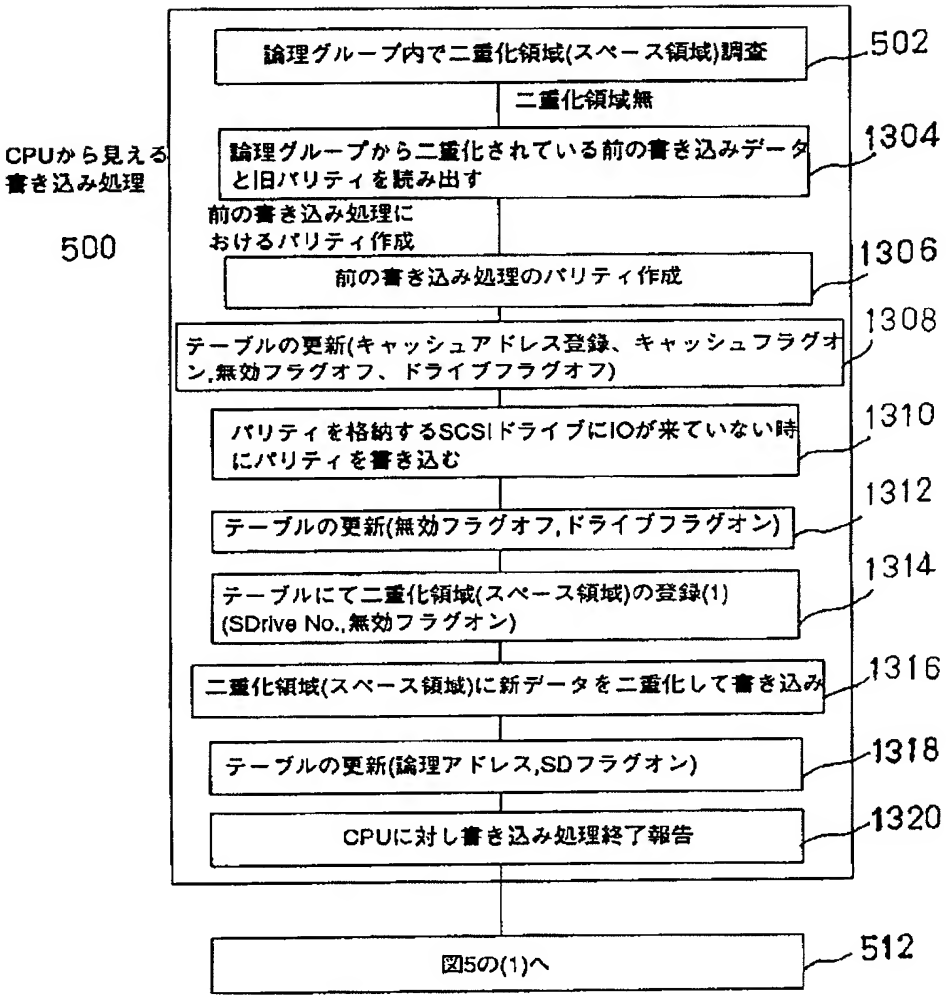
【図１２】

図12


(a)
読みだし回転
待ち0.5回転
書き込み回転待ち1回転
読みだし／
パリティ作成
書き込みデータ／
パリティの書き込み
(b)
インデックス
レコードアドレス
(R)
トラック
インデックス
シリンダアドレス
(CC)
ヘッドアドレス
(HH)
読み書きヘッド
磁気ディスク

【図１３】

図13



(1)二重化領域は書き込み先に指定したアドレスに書き込まれているデータと二重化されているデータのペアの中でSCSIドライブ番号の小さい方とする

【公報種別】特許法第１７条の２の規定による補正の掲載
【部門区分】第６部門第３区分
【発行日】平成１３年２月９日（２００１．２．９）

【公開番号】特開平６－３３２６３２
【公開日】平成６年１２月２日（１９９４．１２．２）
【年通号数】公開特許公報６－３３２７
【出願番号】特願平５－１２５７６６
【国際特許分類第７版】
G06F 3/06 305
301
11/10 320
【ＦＩ】
G06F 3/06 305 C
301 Z
11/10 320 Z

【手続補正書】
【提出日】平成１２年４月１４日（２０００．４．１４）
【手続補正１】
【補正対象書類名】明細書
【補正対象項目名】特許請求の範囲
【補正方法】変更
【補正内容】
【特許請求の範囲】
【請求項１】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブに格納する記憶装置の制御方法において、
データ書き込みに伴う誤り訂正符号の更新を、前記データ書き込み処理よりも遅延させて処理することを特徴とする記憶装置の制御方法。
【請求項２】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置におけるデータの格納方法であって、前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置におけるデータの格納方法において、
データ書き込みに伴う誤り訂正符号の更新を、前記データ書き込み処理よりも遅延させて処理することを特徴とする記憶装置の制御方法。
【請求項３】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブに格納する記憶装置の制御方法において、
データ書き込みに伴う誤り訂正符号の更新を、データの書き込みを要求した上位置装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。
【請求項４】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置におけるデータの格納方法であって、前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置におけるデータの格納方法において、
データ書き込みに伴う誤り訂正符号の更新を、データの書き込みを要求した上位置装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。
【請求項５】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブ

に格納するにおいて、
データ書き込みに伴う誤り訂正符号の更新を、前記制御装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。
【請求項６】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置におけるデータの格納方法であって、前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置におけるデータの格納方法において、
データ書き込みに伴う誤り訂正符号の更新を、前記制御装置の負荷が所定値以下のときに処理することを特徴とする記憶装置の制御方法。
【請求項７】データを格納する複数のドライブ群と、該ドライブ群を管理する制御装置とを備え、書き込みデータを前記ドライブ群に格納し、該各ドライブに格納されているデータにより誤り訂正符号を生成し、この生成した誤り訂正符号を、該誤り訂正符号を生成するのに関与したデータが格納されているドライブとは別のドライブに格納するディスクアレイシステムにおいて、
データ書き込みに伴う誤り訂正符号の更新を、前記データの書き込みとは非同期でかつ所定のタイミング毎に処理することを特徴とするディスクアレイ装置の制御方法。
【請求項８】あるドライブに障害が発生し、障害が発生したドライブ内のデータまたはパリティを回復処理により復元し、この復元したデータまたはパリティを予備の書き込み領域に書き込んだ後、障害が発生したドライブを、正常なドライブに交換し、交換後は、この交換した正常なドライブは全てスペース領域により構成されているとして論理グループを再構成して処理を再開することを特徴とする請求項１記載の記憶装置の制御方法。
【請求項９】あるドライブに障害が発生したことを感知したら、障害が発生したドライブを正常なドライブに交換し、障害が発生したドライブ内のデータまたはパリティを回復処理により復元し、この復元したデータまたはパリティと、障害が発生したドライブ内にあったスペース領域を、交換した正常なドライブに書き込んで再構成して処理を再開することを特徴とする請求項１記載の記憶装置の制御方法。
【請求項１０】最新に書き込まれたデータについては二重化して高信頼とし、書き込み要求があまり発行されないデータについてはパリティにより信頼性を確保するように、信頼性について二段階のレベルを設定することを特徴とする請求項１記載の記憶装置の制御方法。
【請求項１１】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つの論理グループを生成し、該論理グループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する記憶装置であって、前記論理グループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つの論理グループを新たに生成し、前記新たな論理グループの更新データと誤り訂正符号とを、前記書き換え前のデータ及び誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ分散して格納する記憶装置において、データ書き込みに伴う誤り訂正符号の更新を、前記データ書き込み処理よりも遅延させて処理することを特徴とする記憶装置。
【請求項１２】前記記憶装置を構成するドライブの集合の中に、書き込むデータを一旦二重化して書き込める領域をもつことを特徴とする請求項１１記載の記憶装置。
【請求項１３】スペース領域を、パリティを生成するデータと、それらのデータにより生成したパリティが格納されているドライブに分散してもつことを特徴とする請求項１１記載の記憶装置。
【請求項１４】スペース領域を、パリティおよびパリティの作成に関与したデータの格納されているドライブ以外の、異なる２台のドライブに確保することを特徴とする請求項１１記載の記憶装置。
【請求項１５】論理グループにおいて、パリティを作成するデータとパリティの集合をパリティグループとし、書き込み前と書き込み後では、パリティグループを構成するデータおよびパリティの格納されているドライブが異なることを特徴とする請求項１１記載の記憶装置。
【請求項１６】論理グループ内のあるドライブに障害が発生した場合、正常な残りのドライブ内のデータとパリティから、障害が発生したドライブ内のデータまたはパリティを回復処理により復元するが、この復元したデータまたはパリティをスペース領域に書き込む制御を行うプロセッサを持つことを特徴とする請求項１１記載のディスクアレイシステム。
【請求項１７】書き込みにより作成した新パリティを、当該ドライブへ書き込む前に、あるドライブに障害が発生した場合、旧パリティの作成に関与したデータについては、この旧パリティの作成に関与した、正常な残りのドライブ内のデータとパリティから回復処理により復元し、二重化されている新データが格納されているドライブに障害が発生した場合は、二重化データの一方から、障害が発生したドライブ内のデータまたはパリティを回復処理により復元する制御を行うプロセッサを持つことを特徴とする請求項１１記載のディスクアレイシステム。
【請求項１８】書き込みにより作成した新パリティを、

当該ドライブへ書き込む前に、あるドライブに障害が発生した場合、正常な残りのドライブ内のデータとパリティと二重化されている新データから、障害が発生したドライブ内のデータまたはパリティを回復処理により復元するが、この復元したデータまたはパリティをスペース領域に書き込む制御を行うプロセッサを持つことを特徴とする請求項１１記載のディスクアレイシステム。

【請求項１９】外部装置から入力されたデータを記憶装置に書き込む方法であって、

入力データを第１の冗長レベルの形式で上記記憶装置に書き込むステップと、

上記記憶装置に第１の冗長レベルの形式で書き込まれた上記データを第２の冗長レベル形式に書き換えるステップとを有することを特徴とするデータ書き込み方法。

【請求項２０】前記外部装置と前記記憶装置との間のデータの入力または出力の頻度を検出するステップを有し、

上記検出された頻度に応じて、前記第１の冗長レベルの形式で書き込まれたデータを第２の冗長レベルの形式に書き換えることを特徴とする請求項１９記載のデータ書き込み方法。

【請求項２１】前記第１の冗長レベルの形式が２重化を利用するものであり、前期第２の冗長レベルの形式がパリティを利用するものであることを特徴とする請求項１９または請求項２０に記載のデータ書き込み方法。

【請求項２２】外部装置と接続可能であって、上記外部装置からの入力データを書き込み、内部に記録されているデータを上記外部装置に読み出す記憶装置であって、

上記外部装置からの入力データを第１の冗長レベルの形式で書き込む手段と、

上記第１の冗長レベルの形式で書き込まれたデータを第２の冗長レベル形式に書き換える手段とを有することを特徴とする記憶装置。

【請求項２３】前記外部装置と前記記憶装置との間のデータの入力または出力の頻度を検出する手段を有し、

前記書き換え手段が、上記検出手段によって検出された頻度に応じて、前記第１の冗長レベルの形式で書き込まれたデータを第２の冗長レベルの形式に書き換えることを特徴とする請求項２２記載の記憶装置。

【請求項２４】前記第１の冗長レベルの形式が２重化を利用するものであり、前期第２の冗長レベルの形式がパリティを利用するものであることを特徴とする請求項２２または請求項２３に記載の記憶装置。